

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:39 ; Search time 100.46 Seconds

(without alignments)
1384.761 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRTWESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq.101002.*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5216	96.6	1063	AAAM39991	Human polypeptide
2	5165.5	95.7	1050	ABG10171	Novel human diapo
3	3369	62.4	935	ABG10172	Novel human diapo
4	2378	44.1	537	AAAM41777	Human polypeptide
5	1800.5	33.0	1011	ABR71113	Drosophila melanog
6	1782	33.0	700	AAAB64615	Human secreted pro
7	1770.5	32.8	697	AAAB64616	Human secreted pro
8	1738.5	32.2	1026	AAU23125	Novel human enzyme
9	1714.5	31.8	1098	AAU57544	Human Myosin-1F pr
10	1685	31.2	909	AAU23128	Novel human enzyme

11	1565	29.0	1089	23	ABR97258	Novel human protei
12	1412.5	26.2	759	22	ABG20610	Novel human diapo
13	1362	25.2	2167	22	ABR60369	Drosophila melanog
14	1330.5	24.6	512	21	ABR42651	Human ORFX ORF2415
15	1328	24.6	2129	22	ABR62828	Drosophila melanog
16	1301.5	24.1	1493	21	ACG48639	Arabidopsis thalia
17	1301.5	24.1	1544	21	ACG48639	Arabidopsis thalia
18	1294	24.0	1483	21	ACG48640	Arabidopsis thalia
19	1284.5	23.8	2058	23	ABR97219	Novel human protei
20	1272.5	23.6	2424	22	ABR58924	Drosophila melanog
21	1269	23.5	1502	21	ACG30439	Arabidopsis thalia
22	1269	23.5	1536	21	ACG30438	Arabidopsis thalia
23	1261.5	23.4	1495	21	ACG30440	Arabidopsis thalia
24	1241.5	23.0	2048	22	AAEL1891	Angiogenesis assoc
25	1241.5	23.0	2057	22	AAEL1890	Angiogenesis assoc
26	1223.5	22.7	1754	21	ACG52410	Arabidopsis thalia
27	1223.5	22.7	1769	21	ACG52409	Arabidopsis thalia
28	1223.5	22.7	1804	21	ACG52408	Arabidopsis thalia
29	1218	22.6	1839	21	AAV94291	Caenorhabditis ele
30	1216	22.5	764	22	AAAM80123	Human protein SEQ
31	1208	22.4	697	17	AAAM0603	Human non-conventi
32	1191	22.1	753	22	AAAM79139	Human protein SEQ
33	1191	22.1	1939	23	ABR77096	Human alpha-myosin
34	1187.5	22.0	1453	22	AAH39213	Human polypeptide
35	1186	22.0	369	21	ABR57005	Human prostate can
36	1183.5	21.9	1469	22	AAH39214	Human polypeptide
37	1182	21.9	1120	21	AAV94292	Helianthus annuus
38	1181.5	21.9	631	17	AAAM0604	Murine myosin VII
39	1181	21.9	1879	22	AAAM25750	Human protein sequ
40	1174.5	21.8	1960	22	AAH78854	Human protein SEQ
41	1174.5	21.8	1963	22	AAH78854	Human protein SEQ
42	1172.5	21.7	2056	22	ABR59344	Drosophila melanog
43	1171	21.7	1988	22	AAAM40999	Human polypeptide
44	1171	21.7	1988	22	AAAM41000	Human polypeptide
45	1169.5	21.7	1972	17	AAAM00024	Smooth muscle myos

ALIGNMENTS

RESULT 1
AAAM39991 standard; Protein: 1063 AA.

AC AAM39991:
22-OCT-2001 (first entry)
Human polypeptide SEQ ID NO 3136.

DE Human: noctropic; immunosuppressant; cytostatic; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
leukemia.
KW
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
PD
XX
26-JUL-2001.
XX
PF
XX
26-DEC-2000; 2000WO-DS34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB: AA159147.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Example 4; SEQ ID NO 3136; 10078pp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 XX Sequence 1063 AA:
 SQ
 Query Match 96.6%; Score 5216; DB 22; Length 1063;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1003; Conservative 22; Mismatches 13; Indels 0; Gaps 0;
 QY 7 AAGSGVRYTMSALTAARVGVODVLENTSEAFENLRFRRENTLYTYIGPVLY 66
 DB 26 AAGSGVRYTMSALTAARVGVODVLENTSEAFENLRFRRENTLYTYIGPVLY 85
 QY 67 SVNPRDQIYRQHMERVGVSEYVPPHLEFAVDYVYRALRTERDQAVMISGSGAG 126
 DB 86 SVNPRDQIYRQHMERVGVSEYVPPHLEFAVDYVYRALRTERDQAVMISGSGAG 145
 QY 127 KTEATRILOFAETCPAPERGAVBDRLQSNPVLEAFGNKTLRNDSRFKYMVQ 186
 DB 146 KTEATRILOFAETCPAPERGAVBDRLQSNPVLEAFGNKTLRNDSRFKYMVQ 205
 QY 187 FDEKAPVGHILSYLLEKSRVYHONHGERNFHVFQLLGESEETLRRLGLERNPOSTL 246
 DB 206 FDEKAPVGHILSYLLEKSRVYHONHGERNFHVFQLLGESEETLRRLGLERNPOSTL 265
 QY 247 YLVKGCACAVSSINDSKMKVARKALSVDFTEDEVEDLSTIVASLHIGNHFADEDS 306
 DB 266 YLVKGCACAVSSINDSKMKVARKALSVDFTEDEVEDLSTIVASLHIGNHFADEDS 325
 QY 307 NAGVTENOLKUTRLIGVEGTLRREALTHRKIIAAGEELLSPNLNQAAVARDALAKAV 366
 DB 326 NAGVTENOLKUTRLIGVEGTLRREALTHRKIIAAGEELLSPNLNQAAVARDALAKAV 385
 QY 367 YSRFTWLVKIRNSLASKDAESPWSRSTTVGLDLYGEVYVQHNSSFQFCINYNEKL 426
 DB 386 YSRFTWLVKIRNSLASKDAESPWSRSTTVGLDLYGEVYVQHNSSFQFCINYNEKL 445
 QY 427 QQLFIETLTKSEDEEYEAAGIAPEVOYFNKKIICDLVEKFGIISIDEECLRGCEAT 486
 DB 446 QQLFIETLTKSEDEEYEAAGIAPEVOYFNKKIICDLVEKFGIISIDEECLRGCEAT 505
 QY 487 DLTFLEKLEDTVRPHFELTHKADQKTRKSLDRGEFRLHYAGEVYVSVTGLDRNDL 546
 DB 506 DLTFLEKLEDTVRPHFELTHKADQKTRKSLDRGEFRLHYAGEVYVSVTGLDRNDL 565

QY 547 LERNKETMSSNPIAOCDFKSELSDDKREBEVATOFKMSLQVLEILSKEPAYIRC 606
 DB 566 LERNKETMSSNPIAOCDFKSELSDDKREBEVATOFKMSLQVLEILSKEPAYIRC 625
 QY 607 IKPNDAKQGRDEVLIRHQQVYGLMENVLRVRAGFAVRRKYEFQIRYKSLCEWPM 666
 DB 626 IKPNDAKQGRDEVLIRHQQVYGLMENVLRVRAGFAVRRKYEFQIRYKSLCEWPM 685
 QY 667 WAGRPDQAVAVLVRHLGYRPEEKMGRTKIFIRFETPATDSLEVRQSLATKIQAAW 726
 DB 686 WAGRPDQAVAVLVRHLGYRPEEKMGRTKIFIRFETPATDSLEVRQSLATKIQAAW 745
 QY 727 RGFHWROKELRVKRSALCIQSMWRTGLGRKAKAKKAAQTIRRLIRGLHSPRCPE 786
 DB 746 RGFHWROKELRVKRSALCIQSMWRTGLGRKAKAKKAAQTIRRLIRGLHSPRCPE 805
 QY 787 AFLDHYRASFLNLRQPLRNVDLTSWPPPALREASELLRELCKMMVKYCSISP 846
 DB 806 AFLDHYRASFLNLRQPLRNVDLTSWPPPALREASELLRELCKMMVKYCSISP 865
 QY 847 EMKQLOQKAVASEIFKGGKDNTPQSVPRLESTRGTEISPRVLOSGSEPIQYAVV 906
 DB 866 EMKQLOQKAVASEIFKGGKDNTPQSVPRLESTRGTEISPRVLOSGSEPIQYAVV 925
 QY 907 VKYDRKGYRPPQOLITPSAVYVVEDAKYKORIDVANTLGISVSLDSILVHQRBD 966
 DB 926 VKYDRKGYRPPQOLITPSAVYVVEDAKYKORIDVANTLGISVSLDSILVHQRBD 985
 QY 967 NKQGDVVLQSDHYETLFTALSDRVNNININOSITPAGGPGHDIIDFTSGSELL 1026
 DB 986 NKQGDVVLQSDHYETLFTALSDRVNNININOSITPAGGPGHDIIDFTSGSELL 1045
 QY 1027 TKAKNGHLAVVAPRLNSR 1044
 DB 1046 TKAKNGHLAVVAPRLNSR 1063
 RESULT 2
 ID ABG10171 standard; Protein: 1050 AA.
 XX ABG10171;
 AC 13-FEB-2002 (first entry)
 DT
 DE Novel human diagnostic protein #10162.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-0508631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR N-PSDB: AA574358.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostic, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

xx Claim 20: SEQ ID No 40530; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/Published_pct_sequences.
 CC
 xx Sequence 1050 AA;

Query Match 95.7%; Score 5165.5; DB 22; Length 1050;
 Best Local Similarity 95.8%; Pred. No. 0;
 Matches 995; Conservative 24; Mismatches 19; Indels 1; Gaps 1;

QY 7 ALGSDGVATMESALTARDRVGVDFVLENTSEAFIENLRFRRENLIYTYIGPYLV 66
 DB 12 ALGSDGVATVMDSALTARDRVGVDFVLENTSEAFIENLRFRRENLIYTYIGPYLV 71
 QY 67 SVNPRDLQIYSROHMERKGVSEYVPHLEFAVDYVYALRTERRDQAVMISGESGAG 126
 DB 72 SVNPRDLQIYSROHMERKGVSEYVPHLEFAVDYVYALRTERRDQAVMISGESGAG 131
 QY 127 KTEARKRLLOFYAEICPAERGVANRDLLQSNPYLAEFGNAKTLRNDNSSRFKYMVO 186
 DB 132 KTEARKRLLOFYAEICPAERGVANRDLLQSNPYLAEFGNAKTLRNDNSSRFKYMVO 191
 QY 187 FDFKAPYVGHILSYLLEKSRVYHONGERNHFVYQLLEGEEETLRRLLENNPOSYL 246
 DB 192 FDFKAPYVGHILSYLLEKSRVYHONGERNHFVYQLLEGEEETLRRLLENNPOSYL 251
 QY 247 YLVKQCAKAVSSINDSKDMKYRKALSYIDFTEDEVEDLLSIVASVHLGNIHFAADEDS 306
 DB 252 YLVKQCAKAVSSINDSKDMKYRKALSYIDFTEDEVEDLLSIVASVHLGNIHFAADEDS 311
 QY 307 NAOVVTENOLKYLTPFLLEVEGTLTREALTHRIIAKGBELLSPNLBQAARAARALAKAY 366
 DB 312 NAOVVTENOLKYLTPFLLEVEGTLTREALTHRIIAKGBELLSPNLBQAARAARALAKAY 371
 QY 367 YSRFTFWLVRKINRSLASKDAESPSMRSTYVGLLDIYGFEVQHNSEFOFCINVCNEKL 426
 DB 372 YSRFTFWLVRKINRSLASKDAESPSMRSTYVGLLDIYGFEVQHNSEFOFCINVCNEKL 431
 QY 427 QQLFTLTLKSEQEEYEAEGIAVEPVQYFNKKIICDLVEKFKGIISILDEECLRGEAT 486
 DB 432 QQLFTLTLKSEQEEYEAEGIAVEPVQYFNKKIICDLVEKFKGIISILDEECLRGEAT 491
 QY 487 DLTFLKLELDYKPRPHFLTKHLADQKTRKSLDRREFLLHYAGVTVSYVGLDKNDL 546
 DB 492 DLTFLKLELDYKPRPHFLTKHLADQKTRKSLDRREFLLHYAGVTVSYVGLDKNDL 551
 QY 547 LFRNLKETWCSMNPIMACFDKSELSDKKRPETVATQFKMSLDLVELLSKPEPAYTRC 606
 DB 552 LFRNLKETWCSMNPIMACFDKSELSDKKRPETVATQFKMSLDLVELLSKPEPAYTRC 611
 QY 607 IKPNDAKQGRFDEVILRHQVYKGLMENVLRRAAGVYRRKYEAFLOKYSCLPETWPM 666

DB 612 IKPNDAKQGRFDEVILRHQVYKGLMENVLRRAAGVYRRKYEAFLOKYSCLPETWPM 671
 QY 667 WAGRPDDGVAVALVRHLGYKPEEYKMGRTKIFRPEPTLFATEDSLVRROSLATKIOAM 726
 DB 672 WAGRPDDGVAVALVRHLGYKPEEYKMGRTKIFRPEPTLFATEDSLVRROSLATKIOAM 731
 QY 727 RGFHMRQKFLTVRSALICIOSMMWGTGGRKAKRKMAAQTIRLIRGFI LRHSRCPEN 786
 DB 732 RGFHMRQKFLTVRSALICIOSMMWGTGGRKAKRKMAAQTIRLIRGFI LRHSRCPEN 791
 QY 787 AFFLDHVRASELLNLRQLPRNVLDTSWPTPPALREASBELRELCKMMWVKYCRSIS 846
 DB 792 AFFLDHVRASELLNLRQLPRNVLDTSWPTPPALREASBELRELCKMMWVKYCRSIS 851
 QY 847 EMKQOQKRVASEIRKGRKDNVPOSVPRFISTRGTEISPRVLOSISGEPIQAVPV 906
 DB 852 EMKQOQKRVASEIRKGRKDNVPOSVPRFISTRGTEISPRVLOSISGEPIQAVPV 911
 QY 907 VKYDRKGYKRPRLQTLTPS-AVVIYEDAKVKORIDYANLTGISVSLSDSLFVLHVORE 965
 DB 912 VKYDRKGYKRPRLQTLTPS-AVVIYEDAKVKORIDYANLTGISVSLSDSLFVLHVORA 971
 QY 966 DNKQKGVYVQSDHVITETLTKTALSADRVNNININGSTIFAGGPRDGIIDTSSSELL 1025
 DB 972 DNKQKGVYVQSDHVITETLTKTALSADRVNNININGSTIFAGGPRDGIIDTSSSELL 1031
 QY 1026 ITRAKNGHLAVVAPRLNSR 1044
 DB 1032 ITRAKNGHLAVVAPRLNTR 1050

RESULT 3
 ABG10172
 ID ABG10172 standard; Protein; 935 AA.
 XX
 AC ABG10172;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #10163.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 PD
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS74359.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 CC Claim 20: SEQ ID No 40531; 103pp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 935 AA;

Query Match 62.4%; Score 3369; DB 22; Length 935;
 Best Local Similarity 86.0%; Pred. No. 6, 2e-278;
 Matches 662; Conservative 22; Mismatches 20; Indels 66; Gaps 6;

QY 190 KGAPVGGHILSYLLERSRVVHNGERNFVYLLGEGEETLRRLGLERNQSYLYLY 249
 DB 82 KGAPVGGHILSYLLERSRVVHNGERNFVYLLGEGEETLRRLGLERNQSYLYLY 141
 QY 250 KGCAVVSINDSKDMVKKALSVIDFDEVEDLLSYASVHLGNHFADEDSNQ 309
 DB 142 KGCAVVSINDSKDMVKKALSVIDFDEVEDLLSYASVHLGNHFADEDSNQ 201
 QY 310 VTTENOLKYLRLLEGVETLRRLTHRIKIAKEEELSPLEQAAVARDALAKVYSR 369
 DB 202 VTTENOLKYLRLLEGVETLRRLTHRIKIAKEEELSPLEQAAVARDALAKVYSR 261
 QY 370 TFWVLVKINRSIAK---DAESPSR---STTVGLLDLYGEVDFHNSFEQFCINYN 423
 DB 262 TFWVLVKINRSIAKGRPAFPPLSMSCVPTPELMALLPY-----SFEQFCINYN 313
 QY 424 EKLOOLFIELTSEDEYEAEIGIAPPOYFNKTIICLVKEKGIISIDDELRG 483
 DB 314 EKLOOLFIELTSEDEYEAEIGIAPPOYFNKTIICLVKEKGIISIDDELRG 373
 QY 484 EATDLTLEKLEDTVAPHHFLTHKLADQKTRKSLDRGEFRLIHYAGEVYVYGFIDKN 543
 DB 374 EATDLTLEKLEDTVAPHHFLTHKLADQKTRKSLDRGEFRLIHYAGEVYVYGFIDKN 433
 QY 544 NDLLFRNLKETWSSNMPIMAGCFKSELSDKRPPTVATQFKMSLLQVLEILRSKPAY 603
 DB 434 NDLLFRNLKETWSSNMPIMAGCFKSELSDKRPPTVATQFKMSLLQVLEILRSKPAY 493
 QY 604 IRCIKPNDAK-----GGRFDEVIRHOVYVIGLMEVLARVRRAFGAAR 646
 DB 494 IRCIKPNDAK-----GGRFDEVIRHOVYVIGLMEVLARVRRAFGAAR 541
 QY 647 RKEAEFLQRYKSLCPETWPMAGRPDGYAVLVRHLGYRPEEKMGRTKIFTRPKTLFA 706
 DB 542 RKEAEFLQRYKSLCPETWPMAGRPDGYAVLVRHLGYRPEEKMGRTKIFTRPKTLFA 601
 QY 707 TEDSLFVRQSLATKIQAAWRFHMRQKRLRYKRSALICISQWWRGTLGRRAKAKRMAQ 766
 DB 602 TEDSLFVRQSLATKIQAAWRFHMRQKRLRYKRSALICISQWWRGTLGRRAKAKRMAQ 638
 QY 767 TIRRLRGFTLRHSPRCENAFELDHVRSFLLNLRQPRVAVLDTSWTPPALREASE 826
 DB 639 TIRRLRGFTLRHSPRCENAFELDHVRSFLLNLRQPRVAVLDTSWTPPALREASE 698
 QY 827 LIRELCMKNMWKKYCRSISPEWKOLOQKAVASEIFKGGKNDYPOSPRPLFTSTRLGTEE 886
 DB 699 LIRELCMKNMWKKYCRSISPEWKOLOQKAVASEIFKGGKNDYPOSPRPLFTSTRLGTEE 758
 QY 887 ISPRVQLSGSEPIQYAVPVVYKDRKGKPRRQLLTPSAVVIYEDAKV 936

DB 759 ISPRVQLSGSEPIQYAVPVVYKDRKGKPRRQLLTPSAVVIYEDAKV 808

RESULT 4
 ID AAM41777
 AAM41777 standard; Protein; 537 AA.

AC AAM41777;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6708.

Human; nocrotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000MO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

PI WPI: 2001-442253/47.

PI N-PSDB: AAI60933.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6708; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and

XX the encoded polypeptides (AAM38642-AAM42213) with nocrotropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX of the invention may be used to treat diseases of the peripheral nervous

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX localised neuropathies and central nervous system diseases, such as

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: Immune system suppression,

XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX assays for receptor activity, arthritis and inflammation, leukaemias and

XX C.N.S disorders.

XX Note: The sequence data for this patent did not form part of the printed

XX specification.

XX Sequence 537 AA;

Query Match 44.1%; Score 2378; DB 22; Length 537;
 Best Local Similarity 96.7%; Pred. No. 9, 1e-194;

Matches 462; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

OY 7 ALGSGVATMESHALTARDRVGVDFVLENTSEAFIENLRFRRENLIYTYIGPVY 66
 DB 60 ALGSGVATMESHALTARDRVGVDFVLENTSEAFIENLRFRRENLIYTYIGPVY 119
 OY 67 SVNPYRDQIYSRQIMERYRGVSFEVPHLFAVADYRRLARTERRDQAVMISGESGAG 126
 DB 120 SVNPYRDQIYSRQIMERYRGVSFEVPHLFAVADYRRLARTERRDQAVMISGESGAG 179
 OY 127 KTEATKRLQFYAETCPAPERGAVRDLQSNPVLFAFGNAKTLRNDSSRGKYMVQ 186
 DB 180 KTDATKRLQFYAETCPAPERGAVRDLQSNPVLFAFGNAKTLRNDSSRGKYMVQ 239
 OY 187 FDFKAPVGHILSTLLEKSRVAVHNGERNPHVYQOLLEGGEEETLRRLGLENNPOSYL 246
 DB 240 FDFKAPVGHILSTLLEKSRVAVHNGERNPHVYQOLLEGGEEETLRRLGLENNPOSYL 299
 OY 247 YLVKQCAKAVSSINDKSMKVMKALSVIDTEDEVEDLSTIVASVLHGINHFAADEDS 306
 DB 300 YLVKQCAKAVSSINDKSMKVMKALSVIDTEDEVEDLSTIVASVLHGINHFAADEDS 359
 OY 307 NAOVTTENOLKYLRLGVEGTTLRALTRKIIAKGEBLLSPNLQEAAYARDAKAV 366
 DB 360 NAOVTTENOLKYLRLGVEGTTLRALTRKIIAKGEBLLSPNLQEAAYARDAKAV 419
 OY 367 YSRFTTTLVARKINRSLAKDAESPMSRSTVYGLDIYGFYFOHNSFEQICINCKNL 426
 DB 420 YSRFTTTLVARKINRSLAKDAESPMSRSTVYGLDIYGFYFOHNSFEQICINCKNL 479
 OY 427 QOLFELTLKSEQEEYEAEGIAEPVQYFNKKIICDLVEEKFKGIISILDECLRPE 484
 DB 480 QOLFELTLKSEQEEYEAEGIAEPVQYFNKKIICDLVEEKFKGIISILDECLRPE 537

RESULT 5
 ABB71113
 ID ABB71113 standard; Protein; 1011 AA.
 AC ABB71113;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 40131.
 XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 XX 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA
 PI Venter JC, Adams M, Li PMD, Myers EW;
 DR MPI: 2001-656860/75.
 DR N-PSDB: ABL15216.
 XX New Isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Disclosure: SEQ ID NO 40131; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 1011 AA;
 Query Match 33.4%; Score 1800.5; DB 22; Length 1011;
 Best Local Similarity 39.6%; Pred. No. 4,5e-144;
 Matches 426; Conservative 169; Mismatches 341; Indels 139; Gaps 29;

OY 28 GYODVLENTSEAFIENLRFRRENLIYTYIGPVYVNPYRDQIYSRQIMERYRG 87
 DB 8 GYODVLENTSEAFIENLRFRRENLIYTYIGPVYVNPYRDQIYSRQIMERYRG 66
 OY 88 VSFYEPVPHLFAVADYRRLARTERRDQAVMISGESGAGKTEATKRLQFYAETCPAP-- 145
 DB 67 RELFENAPHLFALDASRYVLRKORODTCILISGESGAGKTEATKIMKTIAAVTNQOQ 126
 OY 146 ---ERGAVYRDLQSNPVLFAFGNAKTLRNDSSRGKYMVQDFKAPVGHILSYL 202
 DB 127 NEIR---VKNVLQSNALLETGONAKTNNDSSRGKYMDEDFDKAPVGGIITNVL 183
 OY 203 LEKSRVAVHNGERNPHVYQOLLEGGEEETLRRLGLENNPOSYLVLVKGCAKAVSSINDK 262
 DB 184 LEKSRVAVHNGERNPHVYQOLLEGGEEETLRRLGLENNPOSYLVLVKGCAKAVSSINDK 240
 OY 263 SDKVMKALSVIDTEDEVEDLSTIVASVLHGINHFAADEDSNAOVTTENOLKYLTRL 322
 DB 241 SDVYGTGNAEFTLSESTDEVOITWRTIAVAVHLAGVFEOTIED-ELVYSKQHLKSTAKL 299
 OY 323 LGVEGTTLRALTRKIIAKGEBLLSPNLQEAAYARDAKAVYSRFTTTLVARKINRSL 382
 DB 300 LQVTELTSTALTRVIAAGSNVQKDHNTQAVRGDALAKAIYDRFLMTISRINRAI 359
 OY 383 ASKDAESPSMRSTVYGLDIYGFYFOHNSFEQICINCKNLQOLFELTLKSEBEY 442
 DB 360 LFRSGTQA--RFSNVIGLDYGFELPDSNFEQICINCKNLQOLFELTLKSEBEY 418
 OY 443 EAGIAEPVQYFNKKIICDLVEEKFKGIISILDECLRPEATDLFLKLEDTVPHP 502
 DB 419 QREGIETNITYFNKKIICDLVEOPHGIITAHDEACLSGKVTDDTLGAMOKNLSKHP 478
 OY 503 HFLTKLADOKTRSL-DRGEFLLHYAGEVTVSVTGLDKNDLFR-----NIKET 554
 DB 479 HYSRQL--KPTDKELKHREDFRITHYAGDVYINNGFIEKNKDTLYQDKRLLHNSKDA 536
 OY 555 MCSMNPMLAQCFKSELSDKKRPETVATOKMSLDLVETLSKKEPARYRCLRPNAQO 614
 DB 537 NLSEMPPEGADIRKT---TKRPLTAGTLEFORSMADLVYTLKKEPPYRCLRPNDLKS 592
 OY 615 PGRFDEVYLRHGVYGLMNTLRVRRAGFAVRRKYEAFLORYSCLCETPMW--AGRPD 673
 DB 593 STVDEDERVEHGVYGLMNTLRVRRAGFAVRRKYEAFLORYSCLCETPMW--AGRPD 652
 OY 734 KFLVYKRSALCIGSMWRTGLRRKRAKMAAOTIRLNGEFLRHSRPECNMFIDHV 793
 DB 712 NFKKK-----NATIVRYKAYKLR----- 732
 OY 794 RASFLNLRLRL--PRNVL---TSMTPPALREASSELLR---ELCKMNV--WKYCRS 843
 DB 733 ---SYVQELANLRRAKQKRDGKSIOWPPLAQRKYEAKLHRMFDFWRANMLHKYPRS 790

QY 844 ISPEMKQLOQKAVASEIFKKNKDNYPQSVPLRFLSTRIGTEEISPRVLOSLS----- 897
 Db 791 ---EM-PQLRIQIATATAGLRPPYWGQA--RHWGDIYLANSENGEYKNSIKININ 844
 QY 898 -----EPIDYAVPVKDKRGYRPRQQLITTSAVYVDAK-----VKQRIDYAN 944
 Db 845 HPADGETFOOVLETSFVKKFNHN-KQANPAFVSDSTIHKLDGINKKFKDMRTYKIRE 903
 QY 945 LTGISVSLSDSLFVLVHVEDKQKGDVY--LQSDHVIETLTKLTALSDRVN----- 995
 Db 904 LTGISVSPGRDQLIVH---SSKNK-DLVFSLESEY-----TLKEDRIGEVYGIYC 951
 QY 996 -----NININGSTFFAGGPGGPGIIDFTSGSELLTKRANGHLAVAP 1039
 Db 952 KKYHDLTGELRVNVTNTNISCRDGRARITVEASNVPEPRKGNITFEVP 1006

RESULT 6
 AAB64615
 ID AAB64615 standard; Protein; 700 AA.
 AC AAB64615;
 XX 22-MAR-2001 (first entry)
 DT Human secreted protein BLAST search protein SEQ ID NO: 125.
 DE
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiatic; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 XX MO200077197-A1.
 PN 21-DEC-2000.
 PD 01-JUN-2000; 2000WO-US14934.
 PF 11-JUN-1999; 99US-0138599.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 PX Rosen CA, Ruben SM, Komatsoulis GA;
 PI WPI: 2001-032312/04.
 DR
 XX Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX Disclosure: Page 516-518; 558pp; English.
 PS
 XX The invention relates to the isolation of genes AAF32757-F32803 encoding
 CC the human secreted proteins AAB64549-B64594. The sequence is a search
 CC result from a BLASTX homology search. The genes and proteins are useful
 CC for preventing, ameliorating or treating medical conditions, e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer, and
 CC other cancers of the adrenal gland, bone, bone marrow, breast, and
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, Rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.

SEQ Sequence 700 AA.
 Query Match 33.0%; Score 1782; DB 22; Length 700;
 Best Local Similarity 51.7%; Pred. No. 9,7e-143;
 Matches 358; Conservative 114; Mismatches 203; Indels 18; Gaps 7;

QY 27 VGVQVPLENNTSAAFLIENLRFRFRNLITTYIGPLVSVNYPRLQIYRSRDMERYR 86
 Db 15 IGVGMVLLLEPL-NETETIDNLKKRFDNEIYIIGSVIVNPRSLPTISPEKVEDYR 73
 QY 87 GVSFEVPPHLEFAVADTVYRALRTERDQAVMISGSGAGTEATKRLLOFYAETCPAPE 146
 Db 74 NNNFELSPHIFALSDAAYRSRLRQDDKQCILITIGESAGATEASKLWMSVVAACCKGA 133
 QY 147 RGNARLDLQSNPLPLEAFNAKITLRNNSRFQKYMDVQDFGAPGVGHISLLEKS 206
 Db 134 EYNOYKEQLQSNPLPLEAFNAKITLRNNSRFQKYMDIEPFGDPAGGYSNYLLEKS 193
 QY 207 RYVHONGERNFHFVYOLLEGGEEETLRRLGLENNPQSYLYLVKGOCAYSSINDKDMK 266
 Db 194 RYVQKPRGBERNFHFYQOLLSASEELLYKLLEDPFSRYNTL-SLDSAKANGVDAAANFR 252
 QY 267 VMKRALVIDFTEDEVEDLISVAVLHGINHPAD-----EDSNAQVTTENQKLYLT 320
 Db 253 TVRNAMQIVGFLDHEAEVAVLEVAALVKGNIKFEKPSRVNGLDESKIKKIELEKFAAS 312
 QY 321 RLQVGEETLLEALHKKIILAKGELLPLLEQAAAROLAARAVSRFTWLYRKINR 380
 Db 313 RPASVK-VYLRRAVSFRTVEAKRKRVSTTLNVAQAYAROLAKNLISLFSWLVNRINE 371
 QY 381 SLASDAESPMSRSTTVGLDLYGFYEQHNSFEQPCINCNKELQOLETLELKEQE 440
 Db 372 SIKAQTVKVK-----VGVLDIYGFELFEDNSFEQFINCNKELQOLETLELKEQE 426
 QY 441 EYEAQIAMEPVQYFNNKIIDVYBEKFGIISLIDECIRPGARDTLTEKLEDTVKP 500
 Db 427 EYIRIEDLEMTIDIEFNNAIIDLLENNNTGLAMLDECIURPGVITDETELEKINQVCAT 486
 QY 501 HPFELTH--KLADQKTRKSLDRGEBRLHVAQEVTVSGTGLKNNDLRLNKETPCSS 558
 Db 487 HOFESRMSKCSRLNNTTLPHSCFRIQHTAGKLYOYEGFVDPKNNDLTRDSQAMKKA 546
 QY 559 MNPIMAOCEPKSELS--DKKRPETVAQFKNLSLQVLEILRSKPEAYIRCIKPNDAKOPG 616
 Db 547 DSHLKSLEPFGNPAKVNLRKPPTAGSQFASVATLNRNLQTKNPYIRICIKPNDKKAH 606
 QY 617 RFDEVLLRHQYKTYGLIENLRVRRAGFAIRKRYEAFLOKYSKLCPETPWPMAGRPDQVA 676
 Db 607 IFNESLVCHOIRYIGLLENVRVRAGYAFQOAYEPCLEERYKMLCKQTPHMKGPARGVE 666
 QY 677 VLVRLGVKPEEYKMGRTKIFIRPKTLFATED 709
 Db 667 VLFNELELPVEDEHSFGSKIFIRNPRFLFOLED 699

RESULT 7
 AAB64616
 ID AAB64616 standard; Protein; 697 AA.
 AC AAB64616;
 XX 22-MAR-2001 (first entry)
 DT Human secreted protein BLAST search protein SEQ ID NO: 126.
 DE
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiatic; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 OS Homo sapiens.

Query Match 32.2%; Score 1738.5; DB 22; Length 1026;
 Best Local Similarity 41.6%; Pred. No. 9e-139;
 Matches 404; Conservative 155; Mismatches 339; Indels 73; Gaps 22;

```

QY 28 GVDDFVLENTSEAFIENLRRENTIYIGVLYSVNPDQLQYSQHMERYRG 87
DB 18 GKDPEVLLQOVME-DEMRNLOLREFEGRITYIGVLYSVNPDQLPYGPEAIARYG 76
QY 88 VSEFVPHLEFVAADVVRALRTERDQAVMISGEGSAGKTAIRLLQFYAFETCPAPER 147
DB 77 RELTERPHILYAVANAYKAMKRSRDTIVISGEGSAGKTESKIMQYIAAVTNPSCR 136
QY 148 GGA--VRRLLOSNPVLEAFGNAKTLRDNSSRFKGYMDQDFKAPYGHILSYLER 205
DB 137 AEYERKDVLLKSTCVLEAFGNARTNRNHNSSRFKGYMDINDFPGDPIGHIHSLXK 196
QY 206 SRVYHONGBRNPHVYQVLLLEGGEETLRLGLERPPQSYLYLVGQCAKV---SSI-ND 261
DB 197 SRVYKQHVGRNPHVYQVLLLEGGEETLRLGLERPPQSYLYLVGQCAKV---SSI-ND 261
QY 262 KSDMKVMRKALSVIDFTEDEVDLSTIVASVHLGHNHPAEDSDNAO-----VTTENQL 316
DB 257 EQSHQAVTAMRVIGSPREVEVSVRHLLAIIHLGNIEVEVEEGLOEGGLAVALDEALV 316
QY 317 KLTLLRLLGVEGTTLRALTHRKIIAKGEEELSP-LNLEQAAVARDALAAVSRFTWLV 375
DB 317 DHVAELTATPRDLVRLSLATFVAGSGRELIEKHTAAEASVARDACAAVQRLFEVNV 376
QY 376 RKINSLASKDAESWSRSTVGLDITGFEVPHONSDEOCICVCEKLOQLFIELTL 435
DB 377 NINSVMEPRGRDPRDCKDTIVIGVDIYGFVEFVNSEOCICVCEKLOQLFIELTL 436
QY 436 KSEOEYEYEGIAMEPVQYFNKKICDLYEEKFGIISILDEECLRPGEATDPLFEKLE 495
DB 437 KOEOEYEREGITWGQSEYFNATIVDLVERPHRGLIANDACSSAGTIDRITQLD 496
QY 496 DTVKPHPHLTKLADQTKSLDNG-ERLLHYGEVYTSYTGGLDNNDLFRNLKET 554
DB 497 THRHHLHTSKQLC-PTDKTMEFRGRFRKHVYAGDVYTSYTGGLDNNDLFRNLKET 554
QY 555 MCSNNPIMAOCF-----DKSELSDKKRPETVATQFKMSLLQVLYILSKPEAVYTRCKP 609
DB 555 LVNSTDPTLRAMWPGQDITVEY--KRPLTAGTLEFKSMVALVENLASKPEYKCIKP 612
QY 610 NDAKQPRPDEVLIHQYKYLGLMENLVRRAGVYRRKYEAFLQRYSLCPEWP--MMA 668
DB 613 NEDKVAAGKIDENHCHQYAYLGLLENVVRAGFASRQPSRFLTRYKMTCEYTPMHL 672
QY 669 GRPOGVAVLVRLHGLKPEEYKMGRTKTEIRPKTLFATEDSLVRRQSLATKIOAAMRG 728
DB 673 GSDKAVALLEQHGLO-GDVAFGHSKLFIRSPRLVLEDS-----713
QY 713 FHMROKFLVYKSAICISQSMWGTGLRRKAAKRRKAAOTIRRLTIRGILRSPROENAF 788
DB 714 ---RARLPI--IYLLLOKAWRGTLARMR-CRRLRAIYITIRMR-----RIKVA-----758
QY 759 FLDHVRASFLNLRQLPRNVLDTSWPPPPALRASELLRELCKMKNWYKCSISPEW 848
DB 759 HLAELQRRF--QAAQPELYGRDLYWPLPAVLQPFQDTCALFCRWRARQLVKNIPSD 816
QY 849 KOOLQOKAVSEIFGKKDNY--POSVRLFISIRLIGEEISPVLSL-----GSEP 899
DB 817 MPOIAKAVYAAAGALQGLKQDWCGRAMARDYLSATDNPTASSLFAORLTKLRDKDGEA 876
QY 900 TOYAVPVVYKRGYKPRRQLLTPSAVVVE--DAKYQRIQIDVANTLGISVSLSDS 956
DB 877 VLFSSHVYKVMR-FHKIRNALLLTDQHLKYLDPDQGRVYRAVPLEAVYGLSTVSGDQ 935
QY 957 LFLVLYHQREDN 967
DB 936 LVLVLAHQDD 946

```

RESULT 9
 AAU97544
 ID AAU97544 standard; Protein; 1098 AA.

AAU97544;
 13-AUG-2002 (first entry)

Human Myosin-1F protein MYO1F.

Human; phosphoinositide-binding protein; phosphatidyl acid; MYO1F;
 phosphatidic acid; signal transduction; housekeeping; myosin-1F.

Homo sapiens.

WO200218946-A2.

07-MAR-2002.

23-AUG-2001; 2001WO-GB03791.

23-AUG-2000; 2000GB-0020833.

15-DEC-2000; 2000GB-0030637.

(BABR-) BABRAM INST.
 (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
 (LIMZ/) LIM Z.

Stephens L, Hawkins PT, Holmes AB, Manfava M, Kistakis N;
 Thuring WJF;

WPI: 2002-434908/46.

N-PDSB: ABK52205.

Novel probe for use in assay method for detecting, measuring,
 identifying and/or isolating PA- and/or PIPn-binding protein in a test
 sample, has immobilised phosphatidyl acid attached to solid support -
 Disclosure; Page 134; 163pp; English.

The present invention relates to a new probe comprising or consisting of
 an immobilised phosphatidyl acid derivative attached on to a solid
 support, or a phosphatidic acid (PA) functionalised solid support. The
 probe of the invention is useful in an assay method for identifying
 and/or isolating a protein that binds to the probe. The invention is also
 useful for detecting, measuring, identifying and/or isolating more than
 one type of phosphatidic acid and/or phosphoinositide-binding protein
 (PIPn) from a test sample e.g. a tissue of tissue culture extract. The
 assay is also useful for detecting, measuring, identifying and/or
 isolating phosphatidic acid and/or phosphoinositide-binding proteins in
 a test sample, to detect and/or measure the ability of an agent, applied
 to phosphatidic acid and/or phosphoinositide-binding proteins in
 a test sample, to agonise or antagonise protein-probe binding, and to
 detect and/or measure the ability of an agent, applied to the probe, to
 agonise or antagonise protein-probe binding. The probe is useful for
 identifying an agonist or antagonist of phosphatidic acid/
 phosphoinositide-binding protein-phosphatidic acid/ phosphoinositide
 interaction, and in a single step high throughput screen of candidate
 agonist and/or antagonist. The invention is also useful for identifying
 important proteins for signal transduction, housekeeping and diagnosis.
 The probe is useful as an important research tool in fundamental research
 for diagnostics and drug discovery. The present amino acid sequence
 represents the human myosin-1F protein MYO1F of the invention.

Sequence 1098 AA.

Query Match 31.8%; Score 1714.5; DB 23; Length 1098;
 Best Local Similarity 38.1%; Pred. No. 1.1e-136;
 Matches 400; Conservative 152; Mismatches 304; Indels 193; Gaps 23;

```

QY 28 GVDDFVLENTSEAFIENLRRENTIYIGVLYSVNPDQLQYSQHMERYRG 87
DB 18 GVDVNVLLPQIT-EDAIANLRKREMDYITFYIGVLYSVNPFQMPYFTDREIDLYOG 76

```

QY	88	VSFEVPHLEFVADTVYRALFTRERPROQAVIISSESGAGTEARKRLLOFEALFACPAER	147
Db	77	AAQYENPHIYALTDNMRYRMLNDCEMOCYIISSESGAKTVAAKYMYSKVGGEK	136
QY	148	GGAVRDLLOSNPLEAFGNAKTLRNDSSRFKQYMDYQEDFGAPVGGHLSYLEKSR	207
Db	137	VQHQKIDITLOSNPILLEAFGNAKTVRNNSSRFKQYFEOFGSRGEPGGKISNPLEKSR	196
QY	208	VYHONHEERNPHVYVYOLLEGEETRLRIGLERNPQSTLYVYQCKAKVSSINPKSPMKV	267
Db	197	VYMONNEKERNHITYOLLEGASQORONLGL-MRPDYTYIYNQSDTYQVGDSTDDBRDPGE	255
QY	268	MKRALSYIDFEDEVEDLLSVASVYLHGINHFADEDSN-AQYTTENQKUYLTRLLGVE	326
Db	256	TLSAMOYIGIPPSIQOLVLOVAGILHGINISFC-EDGNARVESYDILLAFPAVYLLGID	313
QY	327	GTLREALTNRKIIIA-----KGELLSPNLMOAAYPADALAKAYVSRFPMYLVAKINRSL	382
Db	314	SGRLQEKILTSKMDSSRNGRGSSEINVTNLNMQAATIRALAKGLYARLFDFLVEALINRM	373
QY	383	ASKDAESPSWSTVILGLIDYGFVEVQHNSEFOFCINYNCEKILQOLFIETLLKSEOEY	442
Db	374	-QKQOEYS-----IGVLDYIGFEIRQKNGFEQFCINFEVNEKILQOLFIETLLKAEQEB	426
QY	443	EABEIANEPVOYFNKKIICDLVEEKKF--GIISLDEBCL--RPGEXTDLFLEKLEBDT	497
Db	427	VQESIRMTPIPIYFNKKVYCCDLLENKLSPPGIMASVLDVCAFMATGGADDTLQAKLOAA	486
QY	498	VKPRPHETLHLDOKRTERKSLDRGEFBLHAYGEVTVSVTGLDKNNDLFRNFKETMCS	557
Db	487	VGTEHEF-----NSWSAG-FVLIHAYGKSVYDSGFCERNRDPVLESDLELMLQT	534
QY	558	SMNPIMAOCPFKSELSDKK-RPEIYADQFKSLDLYELLRSKEPARYRCIKPNDAKPBG	616
Db	535	SEQAFELRMLFPEKLDGDKGRSPAGSIKKQANDLVATLRCSTPHYTRIKTPNKTRAR	594
QY	617	RFDEVLIRHQVYUGLMENLIVRAGAEYARRKYEAFLORYKSLCPETWPMNAGRPODVA	676
Db	595	DWEENRKHQYEVYIGLKENIRVRAGAEYRRQPAKFLQRYAIIIPETWPRMRGDERQVQ	654
QY	677	VLVNHLGKPEEYKMGRTKIPTRPKTIFATEDSLVNRQSLATYIQAAKMGFINRQKFL	736
Db	655	HLRAVNNRPOYQWGSTKTVKKNDESLFLEEYERKFDGFPARTIQAW-----	704
QY	737	RVKRSALICQSWMRGTLGRKRAKKAQAOTIRLIRGFIIRHSPRCENAFFLDHVRAS	796
Db	705	-----RHVYAVKY-----	713
QY	797	FLNLRLROLPRNVLDTSWPTPPRALREASLIRELCMKNMWYKYSRISPEMKOOLQOKA	856
Db	714	-----EEMRE-----	718
QY	857	VASEIFKKQDNYPQSVRLPFSRPLGTGEESIPVYLSIG-SEPIQYAVPVYVYKRGYK	915
Db	719	EASNILNKKERRRNRISNRNFVGYDLGEE-RPELRQFLGKKERNDPDSYTKYRR-FK	776
QY	916	PRPROLLTPSAVVYVEDAK-----VKORIDYANLNGISVSSLSLDSFLVLAHV	962
Db	777	PIKROLITLTPKCVUYVYIGREKKMKKGPEKGOVCEVLKKKVDIALRBSYSLSTRDDDFIL-	834
QY	963	QRENDKQKGDVYLOSDHYEITLTK--ALSADRRNNININGOSTFA-----	1007
Db	835	-QED-----ADSFLESVEKTEFVYSLCKRFEABATRRPLTFESDTLOFRVKKBSW	884
QY	1008	GGPRGDIIDFTSGSELITRAKNGHLAV 1036	
Db	885	GGGGRSVYTFSGFGDLAVLKVGGRTLIV 913	
RESULT 10			
LAU23128			
LAU23128 standard; Protein; 909 AA.			

PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 23-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236527.
PR 29-SEP-2000; 2000US-0236527.
PR 29-SEP-2000; 2000US-0236567.
PR 29-SEP-2000; 2000US-0236568.
PR 29-SEP-2000; 2000US-0236569.
PR 29-SEP-2000; 2000US-0236570.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0246728.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465566/50.
N-PSDB; AAS40998.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neutral, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases

Claim 11; SEQ ID NO 1124; 1180bp; English.

XX The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. CC AAO22915-AAU23814 represent the novel human enzyme polypeptides of the CC invention.
CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcc_sequences.
CC
XX
XX

Sequence 909 AA;

Query Match 31.2%; Score 1685; DB 22; Length 909;
Best Local Similarity 43.7%; Pred. No. 2.7e-134;
Matches 376; Conservative 133; Mismatches 292; Indels 60; Gaps 18;

QY 28 GVDFVLENTSEAFIEMLRRFRRENLYTYIGPLVSVNPRDQIYSRQEMERYRG 87
DB 18 GKPDVLLDQYTM-DEMRNLQLEFKGRITTYIGELVSVNPRDQIYSRQEMERYRG 76
QY 88 VSEFVPPHLEFVADITVYRALKTERROAVMISGSEAGATKTRKLOFYAETCPAPER 147
DB 77 RELYERPPHLYAVANAAVYKAMKRSRDTCTIVISGSEAGATKTRKLOFYAETCPAPER 136
QY 148 GGA--VDRLLQSNPVLEAFGNKTLTNDSSRRGKMDVOPFKGAPVCGHLSYLEK 205
DB 137 AEYERKVDVLLKSTCVLEAFGNATNNHNSRFGKTYMDINPDKGDPICGHHSYLEK 196
QY 206 SRVYHONHGERNFHFYQLLEGGEEETLRIGLERNFOSYLVKGCACAV---Sst-ND 261
DB 197 SRVYKQVGERNFHFYQLLRGSEDKQLHELHERNFAYVNFTHQAGLNNYHSLDSD 236
QY 262 KSDMKVKKALSVYDFEDEVDELSTVAVYLHGNTHFADEDSNAQ-----VTTENQL 316
DB 257 EDSHVAVTEAMRYVGFSEVESVYRITLAIHLGNIETFEVTEGGLOKGLAVAEALV 316

Db 693 W-----RRHVAVRKY-----702
 Qy 786 NAFELDHVASFLLNLRQLPRNVLDTSWTPPALREASLRELCMKMMWKKYCSIS 845
 Db 703 -----EEMKE-----707
 Qy 846 PEMKQLOQKAVASEIFKGGKKNYPQGVPRFLSTRLGTEPISPRVLSIG-SEPIQYAV 904
 Db 708 -----EASNILLNKERRRNSINRFVQDYLGLRE-RELNLQFLGKKERVDFA 755
 Qy 905 PNVYDRGKYPKRRROLLTPPSANVYVEDAK-----VKGRIDYANITGISVS 951
 Db 756 SVKTYDRR-FKPIKRDILTPKCYVIGREKMKGPEKGPVCELLKKKIDIOALRGVSL 814
 Qy 952 SLSDSLFVLHYQREDNKQKGDVYLOSDDVETLTKT---ALSADRVNININIGSTIEA- 1007
 Db 815 TRQDDEFIL---QED-----ADSFLESVKTFFVSLCKRFEATRRPLTFSD 862
 Qy 1008 -----GCPGRDGIIDFTSGSELLITKAKNGLAV 1036
 Db 863 TLQFRVKKEGWGGGGRSVTFESRGFDLAVLKVGGRITLV 902
 RESULT 12
 ABG20610
 ID ABG20610 standard; Protein; 759 AA.
 AC ABG20610;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #20601.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YF;
 DR MPI, 2001-639362/3.
 DR N-PSDB; AAS84797.
 PT
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT biologically for genetic disorders or other traits and to assess
 PT biodiversity.
 PS Claim 20; SEQ ID No 50969; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 759 AA:
 Qy Query Match 26.2%; Score 1412.5; DB 22; Length 759;
 Db Best Local Similarity 38.5%; Pred. No. 3.7e-111;
 Matches 332; Conservative 122; Mismatches 261; Indels 147; Gaps 18;
 Qy 28 GVQDFVLLNFTSEAFIENLRFRRENLIYVIGVLSVNPYRDLQIYSQHMERYG 87
 Db 10 GKDFVLLDQVME-DFVRNLQRFKRIYIYIGEVLSVNPYQDELPGDEALARTQG 68
 Qy 88 VSEYEPRLFAVADTVYRALTERRDQAVMISGESGAGTEATRLQFYAETCPAPER 147
 Db 69 RELYERPHLYAVANAVYKAMKHSRDTICIVISGESGAGTEASKHIMQYIAAVNPSQR 128
 Qy 148 GGA--VRDRLOSNPYLEAFNAKTLRNDSSRPFKYMDOYDFKAPVGHIIISYLEK 205
 Db 129 AEYERKVDVLLKSTCYLEAFNAKTRNNSSRFKRYMDINDFKGPDIIGHIHSYLEK 188
 Qy 206 SRVYHONHGERNFHFYQLLEGEETRLRLGLENNPSYLYLVKGCQAKVSIIDKSDW 265
 Db 189 SRLKOHVGERNFHFYQALD-----SDQSH 215
 Qy 266 KYMKRALSYIDTEDEVEDLSTVASYVLAHLNIFRADEDSNAQ-----VTENQKLYT 320
 Db 216 QAVTEAMRYIGSPVESVHRILIAHLGIEVEVEEGGLQREGVLAEEALYDVYA 275
 Qy 321 RLGLVEGTLRLALHRIKIAKEBLLSP-LNLEQAAVARDALAKAVYSPFTWLVKRN 379
 Db 276 ELTATPRDVLRSLLARTVYASGRELIEKHTAALASARDAKAAVORLEFENVYRNIN 335
 Qy 380 RSLASKDAESPWRSTYVLGLDIYGEFVPOHNSFQPCINVCNEKLOLFTELTJSE 439
 Db 336 SYMEPRGRDPRRDGDTYIGVDIYGEFVFNPSFQPCINVCNEKLOLFTELTJSE 395
 Qy 440 EEEYEGTAMEPVOYENKTIIDIVEEKFGIISLIDECURRGATLTLEKLEDPVK 499
 Db 396 EYEREGITWQSVETFNNAITVDYERPHRGLAVLDACSSAGTITDITFLQTLDMHNR 455
 Qy 500 PAPHFLTHK-----LADQKTRKSLDRG-EFRLLYAGEVTVYVTGFLDKNNDLFRNLK 552
 Db 456 HHLHTYSKOVPRVAVPQWADKTEWFGDRDRKIKRYAGD-----DLK 497
 Qy 553 ETMCSSNMPVIAQC-----DKSELSDKKRPETVATQPKMSLDLQVLEIRSKEPAYIRCI 607
 Db 498 RMLYSTDPPTLRAMPDQODITEYT--KRPLIGTLTKKMSVLAVENTLASKKEFYVCI 555
 Qy 608 KPNDAKQGRFDEVILRHQVYKYLGLMELRYRBRAGFAVRRKYEAFLQYKSLCEETPMW 667
 Db 556 KPNEDKVGAKTDENCRQVAVYGLLEMAVNS----- 587
 Qy 668 AGRPODVAVLVRLGVPRPEYKMGRTKIFTRPRTLFATEDSLVRRQSLATKIQAMR 727
 Db 588 ALKEQHGL-----QGVAVRHSKSLFTRSPRTLTLLRS----- 620
 Qy 728 GFHMOKFLRYKRSALICQSWRKTLGRRKAAKKAQAOTRLRLRGFTLRHSRCPENA 787
 Db 621 -----RRLTPI--IVLLQKAMRGTLAWR--CRRRLAIYITIMFR--RHKVRVRA---- 665
 Qy 788 FFLDHVASFLLNLRQLPRNVLDTSWTPPALREASLRELCMKMMWKKYCSISPE 847
 Db 666 -HLAELQRRF--QAARQPRPLGRDLVWPLPRVAVLPDPPDTCIALFCMRAROLVKNIPPS 722
 Qy 848 WKQOLOQKAVASEIFKGGKKNY 869


```

XX 31-MAR-2000; 2000WO-US08621.
PF
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
PI
PI Shinketsu RA, Leach M;
XX
XX WPI; 2000-602362/57.
DR N-PSDB; AAC76860.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 4015-4016; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antiparkinsonian; neuroprotective;
XX osteopathic; anticonvulsant; antirheitic; immunosuppressive;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
XX antihydrolytic; and antineumatic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SO Sequence 512 AA:
Query Match 24.6%; Score 1330.5; DB 21; Length 512;
Best Local Similarity 53.9%; Pred. No. 26-104;
Matches 260; Conservative 82; Mismatches 129; Indels 11; Gaps 4;
QY 27 VGVDFVLENTSEAFIENLRPRRENLIYTYIGPVLYSVNPRDQIYSRQHMERYR 86
DB 15 IGVDMVLELP-NSETFINLKKRFDHSEITYIGSVIVSNVPRSLPIYSPEKVEEYR 73
QY 87 GVSFVPPHPLFAVADTYVRALRTRERDAVWISGSGAGKEARKRLQFYAECPAPE 146
DB 74 NNFPELSHIFALDEATRSRDODKDCILITSESGKAEKLVMSYAAACGCGA 133
QY 147 RCGAVRDLRLQSNPVLEAFGNKTLRNDSSRFKYMVDQEDFKGAPYGGHILSYLLEKS 206
DB 134 EYNOVKEQLQSNPVLEAFGNKTYRNDSSRFKYMIDEPDKDPLGCVYSNLLLEKS 193
QY 207 RYVHGNHGRNFHVYQQLGEGEETLRRLGIERPQSYLYLVKQCKAVSSINDKSMK 266
DB 194 RYVAKOPGRGRNFHVYQQLGEGEETLRRLGIERPQSYLYLVKQCKAVSSINDKSMK 252
QY 267 VWRKALSVDFTEDEVEDLISIVASVHLGNTHFAADDSN----AQVTENQKLYRL 322
DB 253 TVRNAMQIYGFMDHAESESLAVVAVNLGITEFEPESRVNGLDSEKSKDNELKEICEL 312
QY 323 LGVESTTLREALTHKTIKAGBELISPLNEQAAVARDALAKAVYSRTFTVLVYRKINSL 382
DB 313 TGIQDSVLERAFSFTVEAKQEKVSTLTINVAQAYARADALAKNLYSRLFSLVNRINESI 372

```

```

QY 383 ASKDAESPWRSTVYLGLDIYGFVFOHNSFEQFCINCNKELQOLFELTKSEOEY 442
DB 373 KAOTKVRK-----VGVGLDIYGFELFEDNSFEQFCINCNKELQOLFELTKSEOEY 427
QY 443 EAGTAMEVOYFNNKIITDLYEKFRTIISLDEBCLRGATGLTFLEKLEDTVKPPH 502
DB 428 IREDIEWTHIDYFNNAITDILIENTNGITAMLDECELRPGVTDETLEKLNQVCATHQ 487
QY 503 HF 504
DB 488 HF 489
RESULT 15
ABB62828
ID ABB62828 standard; Protein; 2129 AA.
XX
XX ABB62828;
AC
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 15276.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL06931.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure: SEQ ID NO 15276; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins
XX (AB157737-AB172072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 2129 AA:
Query Match 24.6%; Score 1328; DB 22; Length 2129;
Best Local Similarity 34.7%; Pred. No. 36-103;
Matches 329; Conservative 165; Mismatches 331; Indels 122; Gaps 22;
QY 29 VQDFVLENTSEAFIENLRPRRENLIYTYIGPVLYSVNPRDQIYSRQHMERYR 88
DB 74 VEDMTITGDL-QETITLRNQNRKAKOLITTYGSMVAVNPPQILPITYNRRIOLYRNK 132
QY 89 SFYEVPPHPLFAVADTYVRALRTRERDAVWISGSGAGKEATKRLQFYAECPAPE 148

```

Db 133 SLAELPHIFAISDIAFQRLQRLKENQCVATISGSGAGKTESTKLLIQLA---AISGKH 189
QY 149 GAVRDRLQSNPVLEAFGNAKTLRNDNSRFGKYMDOFDEKAPVGGHILSYLLEKSRV 208
Db 190 SWIEQITIEANPINEAFGNAKTVRNDNSRFGKXIEIRFTPOGAIQGARIQOYLLEKSRI 249
QY 209 VHONGBERNHFVQLEEGEETLRIGL-ERNPQSTLYIVKGCACAKVSSINDKSDMKV 267
Db 250 VFOQSHDERNNYHIFVCMLAGLSTAEERERKLQEOSPSQYHYLAQGGCFLLPGRGAKDFAD 309
QY 268 MRKALSYDFTDEDEEDLSTIVASYLHGNHHPAEDDSN---AQVTENOLKYLTRLG 324
Db 310 IRAAKVLSFKPEEWSLTLALATLHGNLFTATEVANLATAEIDTPTMLQRYAQLG 369
QY 325 VEGTTLREALTHRKIIANGELLSPNLQOAYADALAKAVYSRTFTWLVRKINRSLAS 384
Db 370 IPISALNALTOQRTIFVHGEHYTSLSKKALIEGRDAFVKSLYDQIFVRIVRINETI-N 428
QY 385 KDAESPWSRSTVYGLDLYGFEVYQHNSFQFCINCYNEKLOQLFELTLKSEQDEYEA 444
Db 429 KOVDOP---MNSIGVLDIFGFENDDNSFQOLCINVANENIQQFVGHIFKMEODEXON 484
QY 445 EGIAMEPYOYFNKLTICDVEEKFGIISILDECLRPEATDLPLEKLEDVYKPHPHF 504
Db 485 EHTNMQHIEFQDNOQIIDLIGKPMNMLSLIDESKFP-KGTDTLLEKLH----- 534
QY 505 LTH---KLADOKTRKSLDGEFRLHYAGEVTVSYTGFLDKNDLLFRNKETMCSM 559
Db 535 VOHGNSIYVKGKTQTSL---FGIRHYAGVVMYPLGFLEKNDSFSGLRTRYORST 590
QY 560 NPIMAQCPDKSLSD--KKRPETVATOFKMSLLOVELTRSEKPAYIRICIRPNDAKOPGR 617
Db 591 NKYLVDIRPHENPMDTAKOP-TLCYKFRNSLMDLMTLSQAHYFIRCIKPNKEPEKN 649
QY 618 FDEVLIRHQVYTLGIMENLRVRAGFAVRRKYEAFLQRYKSLCPETWPMAGRPQDGYAV 677
Db 650 FDKELCVRLRYSGMETARIRRAGYPTIRHAYRAVERRLLVPVGPL---EQDCRK 705
QY 678 LVYRLGY---KPEYKMGRTKIFTRFKTLEATED--SLEYRQSLATK---IQAMR 727
Db 706 LARQICEVALPADSDROYGKTKFLR-----DEDDASLELQRSOLMKSIYVTLQRGIR 758
QY 728 GFHMROKFLVRKRSALCIQSMWRGTLGRK-----AAKRWAAQTIR-- 770
Db 759 RYLFRRYMKRYREAITTYQRYWRGRLQRRKYQVMROGPHRLGACIAAQLTTKFTMVROR 818
QY 771 -----LINGFTLRHSPPCENAFPLDHVRASFLNLRQDPRNVLDTSWPTPPALRBA 824
Db 819 TIKLQALSRGYL-----VRKDFOKKLLERRKQNLKKEELIKLAKMKEA 862
QY 825 SELIRELCMKNNVMWYKCRSISPWKQOLOQKAVASPIFGKKDNYPOSVPRLFISTRLGT 884
Db 863 EELLR-----LQQLKEOK-----EREQREQDEKRLQEOQLKA 895
QY 885 EELSPRVLOSLGSEPIQYAVPVYKYDRKGYKPRRQLLTPSAVVIY 931
Db 896 EAAARNALMAAVQOKRRTKPVKQEAAPKAPTLQARNSLPPPTLIV 942

Search completed: July 7, 2003, 14:21:20
Job time : 105.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:19:35 ; Search time 35.4566 seconds

(without alignments)
866.342 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRTWESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1191	22.1	1939	4	US-09-310-187A-1
2	1146	21.2	1120	4	US-09-147-404-1
3	1105	20.5	2548	4	US-09-172-422-1
4	1018	18.9	1886	4	US-08-938-105-3
5	160	3.0	1809	3	US-09-012-515A-12
6	160	3.0	1809	4	US-08-360-144A-12
7	160	3.0	1809	4	US-08-360-144A-12
8	160	3.0	2549	4	US-08-471-112A-3
9	160	3.0	2549	4	US-08-471-112A-3
10	152.5	2.8	103	4	US-08-905-223-395
11	130	2.4	1529	4	US-09-134-001C-3945
12	126	2.3	567	4	US-09-134-001C-3945
13	124	2.3	1093	5	PCR-US93-03077-1
14	123.5	2.3	1333	3	US-09-356-952-2
15	122.5	2.3	1319	2	US-08-290-731C-2
16	122.5	2.3	1336	2	US-08-290-731C-2
17	122	2.3	652	4	US-09-438-833-5
18	122	2.3	813	4	US-09-438-833-5
19	122	2.3	826	1	US-08-785-241-6
20	122	2.3	826	2	US-08-480-473B-2
21	122	2.3	826	3	US-08-480-473B-2
22	122	2.3	826	4	US-08-480-473B-2
23	122	2.3	826	3	US-09-235-217-2
24	122	2.3	826	4	US-09-235-217-2
25	122	2.3	826	4	US-09-438-833-1
26	122	2.3	826	5	PCR-US96-10251-2
27	122	2.3	1657	1	US-08-287-959-1

28	120.5	2.2	874	3	US-08-804-439A-15	Sequence 15, Appl
29	120.5	2.2	874	3	US-08-720-229-15	Sequence 15, Appl
30	120	2.2	2154	4	US-08-841-349-4	Sequence 4, Appl
31	117.5	2.2	1151	4	US-09-134-001C-3242	Sequence 3242, Ap
32	115	2.1	1200	3	US-08-840-006-5	Sequence 5, Appl
33	113	2.1	1151	3	US-08-840-006-5	Sequence 5, Appl
34	112	2.1	810	1	US-08-840-006-6	Sequence 6, Appl
35	111.5	2.1	1181	2	US-08-785-241-7	Sequence 7, Appl
36	111	2.1	1066	4	US-08-488-940-2	Sequence 8, Appl
37	111	2.1	1066	4	US-09-541-782-8	Sequence 8, Appl
38	111	2.1	1194	2	US-08-723-820-8	Sequence 8, Appl
39	110	2.0	787	1	US-08-488-940-1	Sequence 8, Appl
40	110	2.0	1786	4	US-08-574-763-2	Sequence 8, Appl
41	108	2.0	805	4	US-08-973-462-8	Sequence 2, Appl
42	108	2.0	805	3	US-08-480-473B-4	Sequence 8, Appl
43	108	2.0	805	3	US-08-915-213-4	Sequence 4, Appl
44	108	2.0	805	5	US-09-235-217-4	Sequence 4, Appl
45	108	2.0	805	5	PCR-US96-10251-4	Sequence 4, Appl
					US-08-583-276-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-310-187A-1

Sequence 1, Application US/09310187A
Patent No. 6358751

GENERAL INFORMATION:
APPLICANT: Benichou, Gilles

APPLICANT: Fedoseyeva, Eugenia

TITLE OF INVENTION: Involvement of Autoantigens in Cardiac

FILE REFERENCE: UCSF-090

CURRENT APPLICATION NUMBER: US/09/310,187A

CURRENT FILING DATE: 1999-05-12

NUMBER OF SEQ ID NOS: 3

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1939

TYPE: PRT

ORGANISM: Homo sapiens

US-09-310-187A-1

Query Match

Best Local Similarity

Matches 281, Conservative 142, Mismatches 303, Indels 64, Gaps 18;

QY	4	RASALGSDGVRTWME-----SALTARDRY-----GVDFVLENTSEAFIEHLR	49
DB	48	KATILSREGKVIATENGKVTYKEDQVLOQPEPKIDMML-TFLHEPAVLTRNK	106
QY	50	RRPRENLITYIGPVLYSVNPRDQIYSROMEXRYEVPPHLEPAVADTVRAALR	109
DB	107	ERKAMMTYTGCLCVVYVNPVKPLPVNAEVAARKKRSEAPPHIFISIDNAYML	166
QY	110	TERDDAVNISESGAGTEATKRLQFYAETCPAPEG-----GAVRDLLOSNP	160
DB	167	TRENQSTILITESGAGTKVTKRYIYFASIAAGDGKKDNANKKGLIEDIIDANP	226
QY	161	VLEAGNNAKTLLANDSSRGKYMVDYDFKCAPVGHILSYLLEKSRVYVHONHGRFHY	220
DB	227	ALEAAGNNAKTLLANDSSRGKYMVDYDFKCAPVGHILSYLLEKSRVYVHONHGRFHY	220
QY	221	FYQLEGGEEETRLRLGLEPNOSLYLVKQCKVSSINDSKMKYKRAALSYIDTDED	280
DB	287	FYQILSNKKPELIDMLVLTNNPYDAFYSGE-VVASIDSEELMTDGAFAVILGTSE	345
QY	281	EYEDLLSTVAVSLHGNTHFADE-DSNAQVTTENQKYLTRLGVGEGTFLREALTFRKI	339
DB	346	ERAGYVKKLTGAIMHGNMKFKQKOREGDAEDGEDADKSAIYLMGNSADLGLCHPRV	405
QY	340	IAKGELLSPINTDQAAVARADAKAYYSRTFTLVKIRKINSKADAEPSWRSTTVYLG	399

```

1 RESULT 2
2 US-09-147-404-1
3 : Sequence 1, Application US/09147404
4 : Patent No. 6326163
5
6 : GENERAL INFORMATION:
7
8 : APPLICANT: FORSMANN, WOLF-GEORG
9
10 : APPLICANT: RAIDA, MANFRED
11
12 : APPLICANT: BRENNER, BERNHARD
13
14 : APPLICANT: NIER, VOLKER
15
16 : TITLE OF INVENTION: A METHOD FOR THE DIRECT DIAGNOSTIC DETECTION OF
17 : TITLE OF INVENTION: GENETICALLY CAUSED PATHOGENIC POINT MUTATIONS
18
19 : FILE REFERENCE: 10496/P63231US0
20
21 : CURRENT APPLICATION NUMBER: US/09/147,404
22
23 : CURRENT FILING DATE: 1999-04-06
24
25 : NUMBER OF SEQ ID NOS: 3
26
27 : SOFTWARE: PatentIn Ver. 2.1
28
29 : SEQ ID NO 1
30
31 : LENGTH: 1120
32
33 : TYPE: PRT
34
35 : ORGANISM: Homo sapiens
36
37 : FEATURE:
38
39 : NAME/KEY: MOD_RES
40
41 : LOCATION: (26)
42
43 : OTHER INFORMATION: "Xaa" represents Ala or Val
44
45 : NAME/KEY: MOD_RES
46
47 : LOCATION: (59)
48
49 : OTHER INFORMATION: "Xaa" represents Val or Ile
50
51 : NAME/KEY: MOD_RES
52
53 : LOCATION: (143)
54
55 : OTHER INFORMATION: "Xaa" represents Arg or Gln
56
57 : NAME/KEY: MOD_RES
58
59 : LOCATION: (249)
60
61 : OTHER INFORMATION: "Xaa" represents Arg or Gln
62
63 : NAME/KEY: MOD_RES
64
65 : LOCATION: (236)
66
67 : OTHER INFORMATION: "Xaa" represents Gly or Glu
68
69 : NAME/KEY: MOD_RES
70
71 : LOCATION: (403)
72
73 : OTHER INFORMATION: "Xaa" represents Arg, Gln, Leu or Trp
74
75 : NAME/KEY: MOD_RES
76
77 : LOCATION: (453)
78
79 : OTHER INFORMATION: "Xaa" represents Arg or Cys

```

```

NAME/KEY: MOD_RES
LOCATION: (513)
OTHER INFORMATION: "Xaa" represents Phe or Cys
NAME/KEY: MOD_RES
LOCATION: (584)
OTHER INFORMATION: "Xaa" represents Gly or Arg
NAME/KEY: MOD_RES
LOCATION: (587)
OTHER INFORMATION: "Xaa" represents Asp or Val
NAME/KEY: MOD_RES
LOCATION: (602)
OTHER INFORMATION: "Xaa" represents Asn or Ser
NAME/KEY: MOD_RES
LOCATION: (606)
OTHER INFORMATION: "Xaa" represents Val or Met
NAME/KEY: MOD_RES
LOCATION: (615)
OTHER INFORMATION: "Xaa" represents Lys or Asn
NAME/KEY: MOD_RES
LOCATION: (716)
OTHER INFORMATION: "Xaa" represents Gly or Arg
NAME/KEY: MOD_RES
LOCATION: (719)
OTHER INFORMATION: "Xaa" represents Arg or Trp
NAME/KEY: MOD_RES
LOCATION: (723)
OTHER INFORMATION: "Xaa" represents Arg or Cys
NAME/KEY: MOD_RES
LOCATION: (731)
OTHER INFORMATION: "Xaa" represents Pro or Leu
NAME/KEY: MOD_RES
LOCATION: (736)
OTHER INFORMATION: "Xaa" represents Ile or Met
NAME/KEY: MOD_RES
LOCATION: (741)
OTHER INFORMATION: "Xaa" represents Gly, Arg or Tyr
NAME/KEY: MOD_RES
LOCATION: (778)
OTHER INFORMATION: "Xaa" represents Asp or Gly
NAME/KEY: MOD_RES
LOCATION: (797)
OTHER INFORMATION: "Xaa" represents Ala or Thr
NAME/KEY: MOD_RES
LOCATION: (870)
OTHER INFORMATION: "Xaa" represents Arg or His
NAME/KEY: MOD_RES
LOCATION: (908)
OTHER INFORMATION: "Xaa" represents Leu or Val
NAME/KEY: MOD_RES
LOCATION: (924)
OTHER INFORMATION: "Xaa" represents Glu or Lys
NAME/KEY: MOD_RES
LOCATION: (935)
OTHER INFORMATION: "Xaa" represents Glu or Lys
NAME/KEY: MOD_RES
LOCATION: (949)
OTHER INFORMATION: "Xaa" represents Glu or Lys
OTHER INFORMATION: "Xaa" represents Glu or Lys
US-09-147-404-1
Query Match          21.2%; Score 1146; DB 4; Length 1120;
Best Local Similarity 34.9%; Pred. No. 4,le=100; Indels 64; Gaps 19;
Matches 275; Conservative 140; Mismatches 310;
4 RASALGSDGRVME-----SALTARDRY-----GVDFVLLENTSEAFTEINR 49
      :|::||| |         :|||       ||| | | | | | | | | |
Db   48 KAKIVSEGGKATAEIEYQKTGYVEDDYMGNPKRPDIEMAML-TFLHEPAVLYNLR 106
      ::||| | | | | | | | | | | | | | | | | | | | | |
Qy   50 RRRENTLITTYICPVLYSNPIRDQIYSROMEXKRVSYYEVPPLHFAVTYRAIR 109
      |:|||| | | | | | | | | | | | | | | | | | | | | |
Db   107 DRYSGMIIYYISCLFCTVNPPKMLPVIYPPEVAAYAGKKRRSEAPHTFSISNAYQYML 166
      :|::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Qy   110 TERDDAVMIISGSSEGAKEATRRLLOFYAETCPAPERG-----CAVRDLQSNPY 161
      |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

```

Db 167 TDRNOSIITGESGAGKTVNTRKVIQYFAVIAIAGDRSKDQSPCKGLEDOIIIOANPA 226
QY 162 LEAFNAKTLRNDNSSRFKGYMDVDFKAPGVGHIISYLEKSRVYHONHGERNHYF 221
227 LEAFNAKTVNRDNSSRFKGIKIHGAKXKLASADIEYLLKSRVITQLAERDYHIF 286
QY 222 YOLLEGESEETLRRLGLEBNPOSYLYLVKQCAKYSINDKSDMKVMKALSVIDTEDE 281
287 YOLLSKKRPEDLMILITNNPYDYAFISQGE--TVASIDDAELMATDPAFVLTGPTSE 345
QY 282 VEDLSIVASVLIHGINHRADE--DSNAVTTENOLKYLTRLGVBGTILRRALTRKTI 340
Db 346 KSMKYLGAIMHFGMKRKLKQREQAPDGTTEEDKSAVLMGINSADLGLCHPYXK 405
QY 341 ANGEELSPNLNEQAAYARDALAKAVYSRTFTWLVKIRNSLASKDAESPWSRTVGL 400
Db 406 VGEYVYTKGONVOVYATGALAKAVYEMFMNMTYRINATLETOP-----XOYFICV 459
QY 401 LDYGFVFOHNSFEQFCINYCNEKLOLFIELYLSKEQEEYEAEGIANEPVOYFNKII 460
Db 460 LDYAGFEIFDFNSFEQFCINFNEKLOQFFNNHMFLEQEEYKKEGIEWTFIDXGMDLQA 519
QY 461 C-DLYEKFKGIIISIDECRLRGEATDITFLEKLEDYKPRPHLTKHLDQKTR--KS 517
Db 520 CIDLI-EKPMGIMSLIEECMPF-KATDTEFAKLF-----NHLGSAFQKPRNITG 571
QY 518 LDGEFRLHYAGEVYVYTGFLDKNNDLFRNLKETMCSMNPIAOCF----- 567
Db 572 KPEAFSLHYAKIYVYNIIGWLOKKNKDLXVGLVYOKSSXLSTLFAYYACADAPI 631
QY 568 DKSELSDKRP--ETVATOFKMSLQVLEILRSKEPAYIRCIKPDQAKGFRDEVLIRH 625
Db 632 EKGKGAKKGSSFOFYASALHRENLMKLNLTSTHFEVRCIIPNETKSPGVMDPLVMH 691
QY 626 QYKVLGEMNLVBRAGFAVRRKYEAFLORYKSLCEPMPMAGRPDS--VAIVLIRL 682
Db 692 QLRMCVLEIGIRICRGFNRLIYDXORXYILNPA--IXEGFQXSRKXAEKILSSL 749
QY 683 GYKPEYKMGRTKIFIRPKTFLATEDSLEVAROS-LATKIOANRGFMROKFLRV--K 739
Db 750 DIDHQYKFGHTKF--FKAGLLGLLEBRKXRLRIITRIOAGRGVILAKRMEYKLLER 807
QY 740 RSAICTQSW 748
Db 808 RDSLVIQW 816

RESULT 3
US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chlu, Choi Yling
; APPLICANT: Duhli, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IIX AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; FILE REFERENCE: 200130.442
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-172-422-1

```

```

Query Match 20.5%; Score 1105; DB 4; Length 2548;
Best Local Similarity 25.2%; Pred. 1.4e-95;
Matches 320; Conservative 205; Mismatches 370; Indels 376; Gaps 32;

QY 40 SEAFIENLRFRRENLIYTYIGPLVSVNDYRDQIYSRQMBERYGVSEYVPHLFA 99
Db 158 NEKTLLENLRDRFKEKITYTYVSGIYIYNPFKEPLIYPNKXVYKMDHQLGKREPHIYA 217
QY 100 VADIVYRALRTERDQANMIGESGAGCTEATKLLDQYATCAFERGGA--VRDLLO 157
Db 218 VADAVYHMLQKKNQOCIVISGESGSGKTOSTNLH--HLTALSKGFASGVEQITLG 274
QY 158 SNPLYEAGNAKTLRNDSSRFKGYMDVDFKAPGVGHIISYLEKSRVYHONHGERN 217
Db 275 AGPLYEAGNAKTLRNDSSRFKGFIOVNYDETVIGAVYEKILLESRLVYGENHERN 334
QY 218 FHVYQVLEGESEETLRRLGLEBNPOSYLYV-----VKQCAKYSSTI 259
Db 335 YHVFYLLAGASEDEBSAFHL-KOPEEYHYLNQITKRPRLQSWDQCYDSEPDGFTVEGE 393
QY 260 NDKSDMKVMKALSVIDTEDEVEDLSIVASVLIHGINHRADE--EDSNAVTTENOLK 317
Db 394 DLRHDEKRLQAMEVGFLLPRTROTLESLSAILHLGNTCYKKTYRDSIDICNPEVLP 453
QY 318 YLTRLGVEGTTLREALTRKRIIAKEBELSPNLNEQAAYARDALAKAVYSRTFTWLVK 377
Db 454 IYSELLEVKEMLEALVTRKTYVGEKILIPYLAELAVYRNMASSIVSALFDWIVR 513
QY 378 INRSILA-SKDAESSWSNVTYVGLDIDYGEFVFOHNSFEQFCINYCNEKLOLFIELYTLK 436
Db 514 INHALLNSKDELEHNT--KTLISGVLDIFGFEDYENNSFEQFCINFANERLQHYFNQIHF 571
QY 437 SEOEYEAEGIANEPVOYFNKIIODIYEEKFKIISIDECRLRGEATDITFLEKLED 496
Db 572 LEQETRIEGISMHNIDYIDWTCINLSKRPCTLHLHLDSENP--QATNQTLIDKFK- 629
QY 497 TVKPRPHLTKHLDQKTR--SLDRGEFRLHYAGEVYVYTGFLDKNND-----LL 547
Db 630 -----QHNDNSYIEPRVAMEPAFIKIHAGVYKVKQKFRREKNDHMPDVAL 679
QY 548 FRNLKETMCSM----- 559
Db 680 LRSSKNAFISGMIGIDPAVAFRMAILRAFRFVAVAFREAGKRNIRKTHGDDTAPCATLK 739
QY 560 -----NP-----IMAQ-----PD-- 568
Db 740 SMDSEFLQHPVHORSLLEIQRCKEKEYSTRKNPRLPSDLOGMALLNEKNQHDTEPDA 799
QY 569 -----KSELSD----- 574
Db 800 WNGRTGIRQSRSLSSGTSILDKDGFANSTSSKLEBRAGHILTRKKNKSRPALPKHLEV 859
QY 575 -----KRPETVATOFKMSLQVLEILRSKEPAYIRCIKPN 610
Db 860 NSLNLRLTLQDRITKSLHLHKKKPPSISAOFOASLSKMETLQOAEYFVCKIRSN 919
QY 611 DAKQPRDEVLIHNOYKVLGEMNLVBRAGFAVRRKYEAFLORYKSLCEPMPMAGR 670
Db 920 AEKPLRPSDVLVRLQRLYTGMLTVOIRQSGYSKTSFODFSHFVLLPRNITPSKN 979
QY 671 PDQGVAVLVRHLGYKPEYKMGRTKIFI-----REPRTLFA 706
Db 980 IOD-----FFRKININPDWYGKTMVFLKEGDERHODLLHQVLRITIIILQHMFRVLLC 1035
QY 707 TEDSLEVRQSL-----ATKIOANRGFMROKFL 736
Db 1036 RQHFLLRQASVITIQFRWNYLNQYVDAVOKDAFVMAAALQASRAHLERORYL 1095
QY 737 RVKRSALICIQSMWGTGLRKA-----KRWAA-----QITRLI-----RGITLR 778
Db 1096 ELRAAATVIOQKMDYRRRRHMAALICIQARKKAVRESKROEDOKKIIILQSCRGRRAR 1155
QY 779 HSPRC-----PENAFFLDHVRASFLLNLRQDLPRLVLDTSWPTPPALREASELL 828

```

DB 1156 QRFKALKEORLAETPEVNG--LVNTKIGSLFETQSDSEMEDCSFDRNRKIRAECKSVI 1213
QY 829 REICMKNMVKYCRSISPEMKOOLQAKAVASEIFKGDNDYQSVPRLEFISTRLEGEIS 888
DB 1214 ESNRISRESSVDCLSEPKQOERASQSGVDL---QED-----VLVRER----- 1255
QY 889 PVLQSLGSEPIQYAVPVKRYKRRKPRRDLPLPSAVYVEDAKYKORDYANLGI 948
DB 1256 PSLLEDLHOKKVGRA-----KRESRMRLEQAFISLELTKVR-----SLGSI 1298
QY 949 SVSISDSLFVLHVOREDNKQGDVY---LQSDH-----VIELTKTALSADRYN 995
DB 1299 SPS-----EDRMSTELVPELQSPRGTPESSQSGSLFELSTYEQSKMLE 1345
QY 996 NININGSTIF 1006
DB 1346 SVISDEGLQF 1356

RESULT 4
US-08-938-105-3
; Sequence 3, Application US/08938105
; Patent No. 6353151
; GENERAL INFORMATION:
; APPLICANT: Vikstrom, Karen L.
; TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,105
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Mannel M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3595-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1886 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-938-105-3

Query Match 18.9%; Score 1018; DB 4; Length 1886;
Best local Similarity 32.1%; Pred. No. 1,8e-87;
Matches 253; Conservative 136; Mismatches 286; Indels 114; Gaps 17;

QY 4 BASALGSDGVRYTME-----SALTARDRY-----GVQDFVLENTSAAETENR 49
DB 47 KAKYVREGKVTALFENGKTYTVKEDVMOONPREFKIEMAML-TFLHPAVLYNKK 105
QY 50 RREFRNLITYIGVLVSNPYRDLQIYSROMERYRGVSFEVPHLFAVADTVYRALR 109
DB 106 ERYAMMITYISGLGCVTVNPKWLPYVAEYVAAYRGGKREAPPHIFISIDNAYQYML 165
QY 110 TERRDQAVMISGSGAKTEATKRLQLQFYAETCPAPERG-----GAVDRLLQSNP 160

DB 166 TDRNOSTLITIGEGACKTIVTKRVIQYFASIALIGRSKNDPNANKGLEQIIQANP 225
QY 161 VLEAFNAKTLRNDNSSRFGKYMDVQDFGAPYGHILSYLLEKSVYHQNGENRHHV 220
DB 226 ALFAFNKATYVRNNSSPFGKFIHFAGNGKLASADIEYTLLEKSHVIFQLAERNYHI 285
QY 221 FYOLLEGSEETLRLRIGERNPOSYLYLVGQCAKVSINDKSPMKYRRALSYIDPTE 280
DB 286 FYQLTSNKKPELLMLLTNNPNFYAVSQGE--VSASIDSELLATDSAPFLVGETAE 344
QY 281 EVEDLLSVASVTLHGNHFAADE--DSNAQVTTENOLKYLTRLGVEGTLRLBALHRT 339
DB 345 EKAGVYKLTGAIMHYGNMKFKOKOREQAEDEGTEDADKSAIYMLGNSADLLKGLCHPY 404
QY 340 IAKGELLISPLNQAAYARALAKAYSRFTLVYKIRSLASKRAEBSRSTTVLG 399
DB 405 KVGNERYTKGSOVOOYYSIGALAKSVYEKKFMNVRINATLETQPRQ-----YFIG 458
QY 400 LLDIYGFVFNHNSFEQFCINYCNEKLOQLFIELTKSEQEYEAEGIAPEYQYFNKI 459
DB 459 VLDIAGFE----- 466
QY 460 ICDLVEEKFGIISILDEECIRPGCATDLFLEKLEDTYKPHPHLTKLADOKTR--KS 517
DB 467 ISSLPHKIMGIMSLIEECMF--KADDMFKAKLYD-----NHLGKSNPFQKPRVKG 519
QY 518 LDRGEFLHAYGAVTSYNGEFLDKNDLLFRLNKEIMCSPMNPINACDFKSLSD--- 574
DB 520 KOEAFSLVHYAGTVNITLGLWLEKNKDPLEWYVGLYKSSKLMTLFTSYASADTGD 579
QY 575 -----KKRP---EYVATQFKMSLDLVELIRSEKAPYIRCIKPNDAKPGREDEVILR 624
DB 580 SGKGGKGGKSSQYTSALHRENLTMTLRTTHHFVRCILIPERRKAPGVMDNPLV 639
QY 625 HQVYTLGEMENLVRPRAGFARRRKYEAFLORYKSLCEPTWPMAGROD---GVAVLYRH 661
DB 640 HOLRNGVLEIRICRKGFPNRIILYGDFOYRILNPAALP--EGQFIDSGKAELKLG 697
QY 662 LGVPEEYKMKRTKIFIRFPTLPATDSLEVRROSLATYIOAARGFHROKFLARY--R 739
DB 698 LIDIDHNYKRGHTVYFR--AGLLGLEEMRDERLSRIITRIQANARQOLRIEKKMYER 756
QY 740 RSAICIGSW 748
DB 757 RDALLVIOW 765

RESULT 5
US-09-012-515A-12
; Sequence 12, Application US/09012515A
; Patent No. 6127521
; GENERAL INFORMATION:
; APPLICANT: Berlin, Vivian
; APPLICANT: Chiu, Maria Isabel
; APPLICANT: Cottarel, Guillaume
; APPLICANT: Damagnez, Veronique
; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,515A

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/360,144
 FILING DATE: 20-DEC-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: APV-036.02
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1809 amino acids
 TYPE: amino acid
 TOPOLOGY: 1linear
 MOLECULE TYPE: protein
 US-09-012-515A-12

Query Match 3.0%; Score 160; DB 3; Length 1809;
 Best Local Similarity 19.1%; Pred. No. 2e-05;
 Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;

137 FYAETCPAPERGAVER--DRLLOS-----NPVLEAFGNKTLRN--DNSS--- 177
 382 FDAPAPLPSPSRKALETVDRLTESLDFDYASRIHPIVRLDQSPDLSTAMDTLSLV 441
 178 -REGYMDQFDFKAPVGGHILSYLLEKSRVYHONH-----GERNFHYF 222
 442 FOLGKRYOIFIP-----VWNVVLVRRHNRHQRVYVLCIRIYKGYTLADEEDPLIY 492
 223 Q--LLEGESEETLRRLGLEERNFOSYLVKGCACAKVSSINDKPMVKRAKLSYDFTED 280
 493 QHRMLRSGGDAALASGPVETGPMKLN-----VSTINIQKANGAARR----- 534
 281 EVEDLSTIVASVHLGNHFAADEDSNAOVTTENOLKYLTRLGVEGTTLSREALTHRKII 340
 535 -----VSKDWLEWLRRL----- 547
 341 AKGELLSPNLNQAAVARDALAKAVYSRTFTWLVKRNIRSLASAKDSRSTYVIG- 399
 548 --SLELL-----KDSSPSLRSCWALAQ 568
 400 ----LDIYGEVGFQHNSEFOFCINYCNEKLQOLF--IELTLKSEDEYEAGIAMEP 451
 569 AVNPMARDLF-----NAAFVCSWSELNEDQOELIRISIELALTSQD-----IA-EV 613
 452 VOYFNKKITCDLVEKEFGIISIDE-----ECLAPGEATDITPLEKEIDTVKP 500
 614 TOTLLN--LAERMEHSDKGPLRLDDNGIVLLGERMAKCRAYAKA--LHYKELEFGKGP 668
 501 HPNPL-----THKLADQTRKSLDRGEFRLHYAGEVTVSVTF----- 539
 669 TPALLESLSINNLK--QOPEAAGVLEYAKNH--GELEIATWYKELIHEMEDALVAYDK 725
 540 -LDKNNDLFRNLKETMCSMNPTMA-----QCFDKSELSDKRREPETAQFKMSL 589
 726 KMDTKND-----DEMLIGRMRCLEALGEMQOLHQCCCEKWTLVN--DETOAKMARAA 777
 590 LQVLEILRSKEPVRITRICKPNDANQPGFDEVLLRHQ-----VKYGLMENLRVR 639
 778 AAANGCIGWDSMEETTCMIPRDYHDGAFYRAVLALHDDLESLAOCIDKARDLLDAELTA 837
 640 RAGFAVRRRYEAFLORYKSLCPETWPMAGRPQGVAVLVNHLGYKPEE--YKMGRTKI 696
 838 MAGESYRAYGAMVS--C-----HMSLELEVYQYKL----- 867
 697 FIRPKTLFATEDSLEVRROSLAKIQAMRGFMWROKFLYKRSKAIK-----IQSWRG 752
 868 -----VPERREIIRQIMWRERLOGCCORIVEDQKILMVRSLVSPHEDMTW----- 913
 753 LGRRRAKRWAAQOTIRRLNGFILRHSPRCPENA--FFLDHVRVSFLNLNRQLPRVLD 811

DB 914 -----LKYASLCGKGRALAHKTVLLGVD---PSROLD 946
 QY 812 TSWPTPPALREASELLRELCKMNMWYKCRSISP-----EMKQOLOKAVASEIFK 863
 DB 947 HPLPTVHPQVYTA-----YKMNW-KMSARKIDAFQHMQHFVQTMOQQAHAATDEQ 998
 QY 864 GKKNYPOSVRLTIRLGTSEISPRYLOSSEPI-----QYAVPVYKIDRNGYKPRPR 919
 DB 999 HKOELH-KLMARCEL--KLGEWOLN--LOGINSEITPKVLYQVYSAATEHDRSVYKAMHA 1052
 QY 920 QLLTPSAVV-----IYDAKVKQR-IDVANLGISVSSLSDSFLVHQRGRENKQCD 972
 DB 1053 WAWNFEVLYHKHQONARDEKKLRASGANITNATTAATTAATTAATTSSTEGSSESE 1112
 QY 973 VLIQSD-----HYIETLTKT-----ALSADRVNI 997
 DB 1113 AESTENSPTSPLOKVTEDLSKTLVMTYVPAVGFPFRSISLSKGNL 1160

RESULT 6
 US-08-360-144A-12
 ; Sequence 12, Application US/08360144A
 ; Patent No. 6150137
 ; GENERAL INFORMATION:
 ; APPLICANT: Berlin, Vivian
 ; APPLICANT: Chiu, Maria Isabel
 ; APPLICANT: Cottarel, Guillaume
 ; APPLICANT: Damagnez, Veronique
 ; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: POLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/360,144A
 ; FILING DATE: 20-DEC-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: APV-036.02
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-832-1000
 ; TELEFAX: 617-832-7000
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1809 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: 1linear
 ; MOLECULE TYPE: protein
 ; US-08-360-144A-12
 Query Match 3.0%; Score 160; DB 4; Length 1809;
 Best Local Similarity 19.1%; Pred. No. 2e-05;
 Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;

137 FYAETCPAPERGAVER--DRLLOS-----NPVLEAFGNKTLRN--DNSS--- 177
 382 FDAPAPLPSPSRKALETVDRLTESLDFDYASRIHPIVRLDQSPDLSTAMDTLSLV 441
 178 -REGYMDQFDFKAPVGGHILSYLLEKSRVYHONH-----GERNFHYF 222
 442 FOLGKRYOIFIP-----VWNVVLVRRHNRHQRVYVLCIRIYKGYTLADEEDPLIY 492

OY 864 GKDNYPQSVPLRFLSTRGLTEISPRVLOSIGSEPI-----OYAVPVYKDRGKYPRR 919
 1122 EDAPAPLPSKRAALETVDRLETSLEDFDVASRIIHPVTRLDOSPRLSTAMDTLSLV 1181
 Db 999 HKQELH-KLMARCL-KLGEMOLN---LOGINESTIPKVLQYYSATEHDSWTKAMHA 1052
 OY 920 QLLITPSAVV-----IVEDAKYKOR-IDYANLTGISVSSLSDSLFLVHOREDNKOKD 972
 1182 FOLGKKYQIFLP-----MVNKLVLRRHINQRYVULICRIYKGTTLADEEDPLIY 1232
 Db 1053 WAWMNEAVLHYKHONQARDEKKLRHAGANITNATTAATTAATTTASTEGSSESE 1112
 OY 973 VVLQSD-----HYETFLTK-----ALSADRVNNT 997
 1273 QHRLRSGQDGLASGPVETGPMKRLH-----VSTINLOKAMGAARR----- 1274
 Db 1113 AESTENSPPTSPLOKKTEDSLKTLMTYVPAVOGFFRSISLSRGNNL 1160

RESULT 8
 US-08-471-112A-3
 ; Sequence 3, Application US/08471112A
 ; Patent No. 6313264
 ; GENERAL INFORMATION:
 ; APPLICANT: Molnar-Kimber, Katherine L.
 ; APPLICANT: Fallili, Amedeo F.
 ; APPLICANT: Caggliano, Thomas J.
 ; APPLICANT: Nakanishi, Koji
 ; APPLICANT: Chen, Yanguo
 ; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/471,112A
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/384,524
 ; FILING DATE: 13-FEB-1995
 ; APPLICATION NUMBER: US 08/312,023
 ; FILING DATE: 26-SEP-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/207,975
 ; FILING DATE: 08-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sleiman, Michael T.
 ; REGISTRATION NUMBER: 36,276
 ; REFERENCE/DOCKET NUMBER: 01142, 0058-00000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4000
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2549 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-471-112A-3

Query Match 3.0%; Score 160; DB 4; Length 2549;
 Best Local Similarity 19.1%; Pred. No. 3.7e-05;
 Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;
 OY 137 FYAETCPAPERGAAGR--DRLLQS-----NPLVLEAFGAKTLRN---DNSS--- 177

Db 1122 EDAPAPLPSKRAALETVDRLETSLEDFDVASRIIHPVTRLDOSPRLSTAMDTLSLV 1181
 OY 178 -RFGKYMDVOEDFKGAPVGHILSYLSEKSRVHONH-----GERNFHFY 222
 Db 1182 FOLGKKYQIFLP-----MVNKLVLRRHINQRYVULICRIYKGTTLADEEDPLIY 1232
 OY 223 Q--LLEGEDEETLRRLGLERNPQSTLYLVKQCAKAVSSINDKSPMKVRKALSVIDFED 280
 Db 1233 QHRLRSGQDGLASGPVETGPMKRLH-----VSTINLOKAMGAARR----- 1274
 OY 281 EVEDLSIVASVLIHGNHFADEDSNMQVTTENOLKTYLFLVEGTTLREALTHRKII 340
 Db 1275 -----VSKDMLWLRRL----- 1287
 OY 341 AKGELLSPNLDEQAAYARALANAVYSRTETVLWRKINRSIAKDAESPMSRSTVLG- 399
 Db 1288 --SLELL-----KDSSEPSLRSCMLAQ 1308
 OY 400 -----LDIYGFVYPOHNSFEQFCINCNELQOLF---ITLTKSEDEYEAGIAMEP 451
 Db 1309 AYNPMARDLF-----NAAFVSCWSELNEQODELRSJLSTLALTSOD-----IA-EV 1353
 OY 452 VOYFNKTIICDLVEKEFGIISILDE-----ECLRGBATDLTFLKLEIDVVKP 500
 Db 1354 TOTLLN-LAEFMHSDKGPRLPDDNGLVLLGERAKCRAYAKA---LHKLELFQKP 1408
 OY 501 HPHPL-----THKLADOKTRKSLDRGEFRLHYAGEVTVSVTG----- 539
 Db 1409 TPALLESLSINRKL--QPEAAGVLEYAKHFE-GELEIQAITYEKLHEMEDALVAYDK 1465
 OY 540 -LDKNNDLPRNLKTEKSSMNPIMA-----QCFDSELSDKRPEVATOFKMSL 589
 Db 1466 KMDTNKD-----DELMIGRRCLEALGEMQLOHQQCEKWTLVN---DETQAKMAMAA 1517
 OY 590 LQVLEILRSKEPAYIRCIKPDNAKQGRFDEVILIRHO-----VKYIGLMENLRVR 639
 Db 1518 AAAGLQGMDSMEYTCMIPRDTHDGAFYRAVVALHODLSLMOQCIDKARDLLDAELTA 1577
 OY 640 RAGFAYRRKYBAFLQRRKSLCPETPWMAAGRPDQVAVLVRLHGLKPEE--YKMGRTKI 696
 Db 1578 MAGESYSRAAGAMVS-----C-----HMLSELEVIQYKL----- 1607
 OY 697 FIREPTLFATEDSLEVRROSLATKIOAMRGFMHROKFLRVKRSATC-----IOSMRRGT 752
 Db 1608 -----VPERREIIRQIWMERLOGCQRIVEDMKILMVSLSLVSPHEDMTW----- 1653
 OY 753 LGRKRAAKRWMAOTIRRLIRGLFRLHSPRCPENA-FELDHVRASFLNLRLPRNVLD 811
 Db 1654 -----LKYASLCGSGRLALAHKTLVLLGLVD---PSHQLD 1686
 OY 812 TSWPTPPALREASSELLRELOKNNMYKYGCRISF-----EMKQOLOOKAVASITFK 863
 Db 1687 HPLPTVHPQVITYA-----YKNNM-WKSARKIDAQOMHMFVOTMOOQAOHAIATEDQO 1738
 OY 864 GKDNYPQSVPLRFLSTRGLTEISPRVLOSIGSEPI-----OYAVPVYKDRGKYPRR 919
 Db 1739 HKQELH-KLMARCL-KLGEMOLN---LOGINESTIPKVLQYYSATEHDSWTKAMHA 1792
 OY 920 QLLITPSAVV-----IVEDAKYKOR-IDYANLTGISVSSLSDSLFLVHOREDNKOKD 972
 Db 1793 WAWMNEAVLHYKHONQARDEKKLRHAGANITNATTAATTAATTTASTEGSSESE 1852
 OY 973 VVLQSD-----HYETFLTK-----ALSADRVNNT 997
 Db 1853 AESTENSPPTSPLOKKTEDSLKTLMTYVPAVOGFFRSISLSRGNNL 1900

RESULT 9
 PCT-US95-06722-12
 ; Sequence 12, Application PC/TUS9506722
 ; GENERAL INFORMATION:
 ; APPLICANT:

```

TITLE OF INVENTION: Immunosuppressant Target Proteins
NUMBER OF SEQUENCES: 25
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06722
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 27-MAY-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 20-DEC-1994
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06722-12

```

```

Query Match 3.0%; Score 160; DB 5; Length 2549;
Best Local Similarity 19.1%; Pred. No. 3.7e-05;
Matches 193; Conservative 121; Mismatches 318; Indels 376; Gaps 46;

```

```

137 FYAETCPAPERGAVER-DRLIQS-----NPVLEAFGNKTKLRN--DNSS--- 177
1122 FDAPERPAPSRKALETVDRLTESLDFDTYASRIIHPYRTLDQPELSTAMDITSLV 1181
178 -RRGYMDYQFPFKGAPVCGHLSLLEKSRVYHQN-----GERNFHY 222
1182 FOLGKKYQIFIP-----WNVKVLVHRHRIHORYDVLICRIYKGYTLADEEDPLY 1232
223 Q-LLEGGEETLRRLERNPQSYLYIKGQCAKVSINDKDMKVMKALVIDETED 280
1233 QHRMLSGGDALASPVEFGPKLKH-----VSTINLCKANCAAR----- 1274
281 EWEELLISVAVLHGNHIFPADEDSNAOVTTENDIKYITRLGVEGTTLRREALHRRKI 340
1275 -----VSRDDMLFWLRRL----- 1287
341 AKGEELLSPLEQAAAYARDALAKAYSTFTFWLVAKIRSLASQAESPSMRSTVVG- 399
1288 --SLELL-----KDSSEPSLSQWALQ 1308
400 -----LLDIYGEFVFOHNSFEQCIINYCNELQOLF--TELTKSEGEYEAEGIAWEP 451
1309 AYNPARDLF-----NAFVSCWSELMDQDELIRSELALTSOD-----IA-EV 1353
452 VOYFNKTIQDLVEEFGKIIISLDE-----ECLRGGEATDLTFLKLEDTYVP 500
1354 TQTLN--LAEFMEHSDKGPLRLDNGITVLGERAKARAVAKA--LAYKELEFGKRP 1408
501 HPEFL-----THKLADOKTRKSLDRGEFRLHYAGEVYVYVTF----- 539
1409 TPAILSLISINKL--QOPEAAAGVLEAMKHF--GELTIQATWYKLEHEDALVAYIK 1465
540 -LDKNDLLEFRNKETKSSMNPIMA-----OCFDSKSELSDKKREPVAATQFKMSL 589
1466 KMDTKND--DPEMLGMRCLFALGEMGOLHQCCCKWTLVN---DETQAKMARMA 1517
590 LQVEILRSKEPPIYICIRPNDAKQGRDEVILRQ-----VYLLGMEWLRAR 639
1518 AAAMGIGQWDSMEETCMIPRDTHDGAFAVRAVLHODLFSLAQOCIDARDDLDELTA 1577
640 RAGFAVRRKYEAFLQRYKSLCPETPMWAGRPQDGAVALVRHLGYRPEE---YKMGRTKI 696
1578 MAGESISRAVGAIVS---C-----IHLSTLEEVYQYKL----- 1607

```

```

QY 697 FIREPKTLFATEDSELEVRQSLATKIQAAWNGFHMROKFLVRKSAIC-----IOSWMRGT 752
DB 1608 -----VPEREIRIOWMERLOGCQRIYEDMQKILMVRSLVSPHEMRKW----- 1653
QY 753 LGRRRAKRRKWAQTIIRLLIRGFLIRHSPCPENA--FLDHYASFLINRRLPRNVLD 811
DB 1654 -----LKYASICGSGRLALAHKTLVLLGVD---PSROLD 1686
QY 812 TSWPTPPALREASELLRELOCMNMWVKYCRSISP-----EMKQOLOQKAVASEIRK 863
DB 1687 HPLPTVHPQVYTA-----YKMM-WKSPARKIDATQOHQVQTMQQAQAHAIATEDQ 1738
QY 864 GKRDNPQSVPRLEFTRIGTEISPRVLOSLEPI---QYAVPYKDYRKGYPRRP 919
DB 1739 HKQELH-KIWARCL-KLEGWOLN---LOGINSETPKYLYQYSAATEHDSRYKAWHA 1792
QY 920 QLLITPSAVV-----IVEDAKYQR-IDYANLTGISVSSLSLYLVHQREDNKQGD 972
DB 1793 WAVNFEAVLVHYKHQNOARDEKKRLRHASGANITNATTAATTAATTTASTSGSSESE 1852
QY 973 VVLQSD-----HYLETLTK-----ALSADRVNNI 997
DB 1853 AESTENSPSPLOKKVTEDELSTLMTVTPAVOGEFRSISLSRGNNL 1900

```

```

RESULT 10
US-08-905-223-395
Sequence 395 Application US/08905223
Patent No. 6222029
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-8550
INFORMATION FOR SEQ ID NO: 395:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig-peptide
LOCATION: -72...-1
IDENTIFICATION METHOD: Von Heljne matrix
OTHER INFORMATION: score 3.6
OTHER INFORMATION: seq CTSLLQLYDAISN/EM

```

US-08-905-223-395

Query Match 2.8%; Score 152.5; DB 4; Length 103;
Best Local Similarity 29.2%; Pred. No. 7.1e-07;
Matches 35; Conservative 25; Mismatches 16; Indels 43; Gaps 4;

QY 615 PGARDEVILRHQVYGLIMENLVRRAGFAVRRKYEAFLORYKSLCPETWPMAGRPOG 674
D 2 PDQFDQAVVNLQRLYSQMLETVRLRKAGVARRPQDFYKRYK----- 44
QY 675 VAVLHRLGKYKPE-----EYKMGRTKFTFPPTLATEDSLEVRQ 716
D 45 --VLMRMLAL-PEVVRGKCTSLLOLYDASNSEWOLKTKVFLR-----ESLEQKLEKRR 96

RESULT 11

US-09-134-001C-3945
; Sequence 3945; Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3945
; LENGTH: 1529
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3945

Query Match 2.4%; Score 130; DB 4; Length 1529;
Best Local Similarity 18.2%; Pred. No. 0.011;
Matches 131; Conservative 100; Mismatches 213; Indels 276; Gaps 36;

QY 25 DRVGVODFVLENETSEAF-----IENLRFRFRNLI-----YTGIPVLSVNPYRDLOIY 77
D 49 DMOGIGFVAMNDKNSHDIIEKSLERLRLDRGGVADGIGTGAGINTETIPY---OLF 105
QY 78 SNOHMERVGSFYEVPRPLFAVADTVYRALRTERDDQAVMISGESGAKTEA---TKRL 134
D 106 EOLTEKRYGEGYAV--GLFPSKEKVRDSIHEMFNOYFESEGRKVIYGRVYDITAI 163
QY 135 LQFVATECPAPER-----GCAVADRLLQSNPVLEAFGNMKT----- 170
D 164 AOHVADTMPYIOQVFDITGVKEVEKRLFLARKQIEKSETOQSIDLYFTSLSHRTIYKG 223
QY 171 -LRANDNSRFKYMDOVEFKGAPGHILSYLLEKSRVYHONHERNP-----HYVQ 223
D 224 WLRSDQIK--GLYDLQNEAVOSKIG-----LVHSRSTNTFPSPKRAHPRM 269
QY 224 LLEGGEETLRLRLGERRNQSIVLYLKQCAKAVSSINDKSDKVRKALSYIDFTEDEVE 283
D 270 LHMNGE-----INTIKGNVMMRRARQKIVETLEDEDK 303
QY 284 DLSLTVASVLIHGNTHFAADESNQVTTENQKILITLLGVEGTTLEALTHRKTIANG 343
D 304 -----KHFVYDEGSDSSIVDNLAEFLSLAMEPE-----KA 335
QY 344 BELSLPLNEQAAV--ARDALAKAVYSRTFTWLVKINRSLASKDAESPMSRSTYVGLL 401
D 336 AMLILP---EPMLVNESNDKVRSPY--EFGYSLM-----PMDGPTM----- 373
QY 402 DLYGEVFOHNSFEQFCINYN--EKLOQLFTELFLKSEQEEYEAEGIAMPEVQY----- 454
D 374 -----ISFCNGDKIGAL-----TRNGL--RGRGTYITMDN 402

QY 455 ----FNKKI-ICDLVEEK--FKGIIISLDECELRPGEDATLFE-----KLEDTVK-- 499
D 403 FTVESSEGVIVPENNFAFG-----QINPGKLLVDFLQKRVENNELKTNIANE 454
QY 500 -PHPHFLTKLADQTRKSLDRGEPRLHYAGEVYVSTGFLDKNDLLEFRLK----- 552
D 455 LVEEQW-----LADYKKNKNDLN-----IYQSS-----DWDQTLFRLOKQFAYTK 496
QY 553 -----ETMCSNMPIMAOQFDK--SELSDKRREPVAOQFKMSLLQVEILRSKPA 602
D 497 EDINKYMDLVYINKKDPDIPGAMGYDAPIVANDK--PESLFNFYKQFLAQVYN----- 546
QY 603 YIRCKIPNDAPGPFDEVILRHQVYGLIMENLVRRAGFAVRRKYEAFLORYKSLCP 662
D 547 -----PPIDA-----YREKIVYSELISGSEGNL-----LCPD 574

RESULT 12

US-09-134-001C-3762
; Sequence 3762; Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3762
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3762

Query Match 2.3%; Score 126; DB 4; Length 567;
Best Local Similarity 18.4%; Pred. No. 0.0048;
Matches 95; Conservative 106; Mismatches 172; Indels 142; Gaps 25;

QY 204 EKSRRVHONHERNP-----HVFQLEGGEEETLRLGLERNPQ----- 243
D 60 KKKETTNANHEKNIILIEKDKKIYS--SGINKIDRKHIDNEANPSKLINKNTKGM 117
QY 244 --SYLYLVKGGCAVSSINDKSDKVRKALSYIDFTEDEVEDLSIVASVLIHGNHFA 301
D 118 RYTYKNTIDDKTIVYISGINNE-----IIDQKD-LMKYLSIV-GIVLFTYVYLA 164
QY 302 ADEDSNQVTTENQKILITLLGVEGTTLEALTHRKTIANGBELSLPLNEQAAVARDA 361
D 165 SRSINRYIRIPINEVYATSLA-----DGYVHNV-----PESVYKE----- 202
QY 362 LAKAVYSRTFTWLVKINRSLASKDAESPMSRST-----TVGLDLYGEVFOHNSFE 415
D 203 -TRALFTYT--NDLARRQKLNNSOKIOSNRLKTLLENIPSSVL-MDKHGEIVANNAVY 259
QY 416 QF-----CINVCNEKLOQLFTELTKSEQEEYEAEGIAMPEV--QYFNKKITCD 462
D 260 QVFNPDQMVENKSYIGFIDISIEKLITE--SFREKVIYEDQLEVAINNVTKKYDVSQIP 318
QY 463 LVEEK--FKGIIISLDECELRPGATDPTFLEKLEDVYKHPHFLTHKLADQTRKSLDR 520
D 319 LTKSKKNLQGVVVLH-----DITMOKLEMLRREFVANVSHELTPYT----- 362
QY 521 GEERLHYAGEVYVSTGFLD-----KNNDL--LFRN--LKETMCSNMPIMAOQFDK 569
D 363 -----SINGFAETLIEGAKKNDQSLDPMNLIIIES--NRRESLYTDLDL 406
QY 570 SELSDKRPETVATQFMKSLDLQVEILRSKREPAYIRICIRKNDAPKQREDEVILRHQVY 629
D 570 -----SYLYLVKGGCAVSSINDKSDKVRKALSYIDFTEDEVEDLSIVASVLIHGNHFA 301

LENGTH: 1333
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-893-952-2

Query Match 2.38; Score 123.5; DB 3; Length 1333;
 Best Local Similarity 18.3%; Pred. No. 0.037;
 Matches 212; Conservative 175; Mismatches 369; Indels 405; Gaps 61.

```

OY 70 PYNDLQIYSQHMERKRGVSFYVPPHLEAVADTVYRALRTERDQAVMISGSGAGKTE 129
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 PY---EFSSEBNAPKRWGL---LVPALKKVOGQVHPRL--ESNDALQYVEE----- 50

OY 130 ATRRLQFYAETCPA--PERGAVARLRLS--NPVLE--AFGAKTLRNDSSRFCKYADV 185
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 ---LIQLNLMLCOADPRASDVEEVQKSFPHIDKMAIADQSAIEKRRKRLNLPV 107

OY 186 QDPFKGAPVGGHILSYLLEKSRVYVQNHGERNHVYOLLGEGEETLRRLGLERNPOSY 245
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 E---KIHPLKEVLGYKIDHVSY-----IVALEVISADILKVG-----NY 148

OY 246 L-----YLVKGCACAKSSINDSKVMKALSVDFTEDEVEDLSTIVASVHLGNTHF 300
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 VRIIRHYETIKODIKVACADK-----VLDMFHODVEDI-----NITS 187

OY 301 AADED-----SNAQVTENQDKYL-- 319
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 LTFDEPSTSGEQTYVDYKAFMAEIROYIRELNLIKFRPEFVSNSLFSANDENTFS 247

OY 320 -----TRLLG--VEGTT--LREALTHRIKANGELLSPLNLE--QAAVARDALAKAV 366
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 RIYDIHELKVLGHIEDTVEMTDEGSPHVLGSCFEDLAEBLARDPEVSAARDILRGE 307

OY 367 YSRFTWLVLR-----KINSLS--ASKDA 387
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 HDREFLSQLSKPRALYLKISGEGEAVOYVLPRLLPVYHCHLHFEELLQLEKSEDO 367

OY 388 ESDPSRSTVLCILDY--GFEVYQNSF-----EOPCINCYNE--KIQOLFIEITLSEQ 439
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 EKECEKLOAITALLVNOGSMERICKSLAKRRLSASACRFSSQMKGQOLAK--KME 424

OY 440 EEEVAGIMEPVQYFNKKIICD---LVEEKFQIISILD--BECLA-----PGEA 485
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 IOKNIDMGKIDIGCCNEFIWEGTLFVVGAKHHEHIFLFGMLICKSNHGQRLGAS 484

OY 486 T-----DLTFLEKLE---DTVKRPHPLTHKLADOKTRSLDRGEFRLHYAGEVYS 535
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 485 NAEYRLKEKEFWKRVQINDKODTNEYKHAFFELIKDENS-----VIFS 527

OY 536 VGFGLDKNN-----DLLEFRLKETMSSMNPIMACCFDSELSDDKRPETVAQFMSL 589
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 528 AKSABEKKNNMAALISLQYRSTLERM-----LDVYMLQEEKE-----EQMR 568

OY 590 LQLVLEILSKREPAVYIRCIKPNDAKOPGFDE-----VILRHQVYKLGIMENIRVRAG 642
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 569 LPSADVYFAER-----DSEENIIFEEENMOPKAGIPITAGVYIKIIEIRLYHM-- 617

OY 643 FAYRRKYAFLQRYKSLCPETWPMAGRPDOGVAVLVRLHY--KPEYKMGRTKIFIRFP 701
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 618 YADPFVYRTFLTYRSEC-----KQPELISLIEREIREPERETEDR----- 660

OY 702 KTLFATED-----SLEVR-----QSLATKIQAMNGEFLW--ROKELRVKRSACIGQSW- 748
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 661 ---IAIENGDOPLSAELKRFREKEYIQVQLRVLVNCR--HMVGHNFDEFERDAYLLQRM 715

OY 749 -WRGTLGRKAARAKMAQITRILIR-----GFLHSPRCPEMNAFFLD--H 792
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 716 EFIGV--RGKAMKM--VESTIKIQKKIARONGRPHNTTFOSSP--PIVENHISRPGH 770

OY 793 VRASFLNLR-----ROL-----PRNVLDTSWP-----TPPALR----- 822
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 771 IETPDLTLHPIELARQLLLESPLVNAVQPSSELVGSVMTKEDKEINSPLLMKIMHTN 830

```

```

OY 823 -----EASELRELCMKNMWYKCRSI--SPEMK----- 849
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 831 LTLWEKCIYETENLEERNAVVRITIELLOVQELNNNGVLEEVASAMSSPVYRLDHTF 890

OY 850 QQL--QOKAVASEIFKGNKY-----POSVP--RUEISTRLGTEISPRVLQS 894
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 891 EQIPSRQKKILEBAHEISEDBHKRYKLAKRSINPCVPFEGYLTNLIKTEEGNPEVLKR 950

OY 895 LGSEPIQYA-----VPVYKDRKGKRPRLQOLLTPSAVYIVDAKVKQRIYANLTGISV 950
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 951 HGEELINFKRRKVAETGEGIQYQNPYCLRV-----ESDIKREFLENLPMGNSM 1001

OY 951 -SSLSDFLVHQRDNOK 970
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1002 EKEFTDYLFKSKLEIEPRNK 1022

```

RESULT 15

```

US-08-290-731C-2
Sequence 2, Application US/08290731C
Patent No. 5843646
GENERAL INFORMATION:
APPLICANT: BOWTELL, David Douglas Lawrence
TITLE OF INVENTION: DNA MOLECULES ENCODING MURINE
TITLE OF INVENTION: SON OF SEVENLESS (MSOS) GENE,
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,731C
FILING DATE: 17-OCT-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00068
FILING DATE: 17-FEB-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PL0921/92
FILING DATE: 17-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: Q-36066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1319 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-731C-2

```

Query Match 2.38; Score 122.5; DB 2; Length 1319;
 Best Local Similarity 18.5%; Pred. No. 0.045;
 Matches 217; Conservative 170; Mismatches 359; Indels 425; Gaps 63;

```

OY 70 PYNDLQIYSQHMERKRGVSFYVPPHLEAVADTVYRALRTERDQAVMISGSGAGKTE 129
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 PY---EFSSEBNAPKRWGL---LVPALKKVOGQVHPRL--ESNDALQYVEE----- 50

```

QY 130 ATKRLLOFYAETCPA--PERGAVDRILLOS--NPYLE-ARFNAKTLINDNSRFGKMYD 185
 Db 51 ---LILQILNMLCOQPRASADVEERYOKSPFPHIDKMAIDAOASALEKKRRR----- 100
 QY 186 QFDKRGAPVGHILSTYLLKESRYVHONGHGRNHFVFO-----LLEGEETEIRRLGL 238
 Db 101 -----NPLS--LPAERIHILLREVIGYKIDQVSYIYAVLEYISADILKLYG- 146
 QY 239 ERNPOSYL-----YLKGOCAKVSINDKSDWKVRRKALSYIDETEDVLDLSTYASVL 293
 Db 147 -----NYRNRIRHYEITRKODIKVAMCADK-----VLMDFHODVEDI----- 183
 QY 294 HLGNIHRADED-----SNAOVTEN 314
 Db 184 ---NILSLTDEPSTSGEQTYYDLVKAFMAEIROYLRELNLIIFKREPEVYSNKSLEFSN 240
 QY 315 QUKYL-----FRILG-VEGTT--LREALTHRKIIAKGELLSPINLE-OAAVAR 359
 Db 241 DVENIFSRITVIDIHELSTVLLGHIEDTYEMTDEGSPHPLVSGCFEDLAEELAFDPESTAR 300
 QY 360 DAL-----AKAVYSRT-----FTWLVRKI 378
 Db 301 DILRFGFHGHEFLSOLSKPGALYLOSGEGFKEAVOYVLPRLILAPYHCLHYFELLKOL 360
 QY 379 NRSLASKDASPSNRSTVGLGLDIY--GEVEFOHNSF-----EFCINYNCE-KLOOLF 430
 Db 361 BEK--SEDOEDKECKKQAITALLINQSGMEKICKSLAKRRLSESACRFYSQOKMKQOLA 418
 QY 431 IELTLKSEOEYEAGIAMEPVQYFNKIIDC---LYEEKFKGIISILD--EBCLR--- 481
 Db 419 IK---KMEIOLKINIDGMEGKDIGOCNEFIMEGTLRVGAKHERHIFLEDGIMICCKSNH 475
 QY 482 ---PGFAT-----DLFELEKLEDIVKPPHFLTHKLADQKTRKSLDRGEFR---L 525
 Db 476 GQPLRPGASSAYRLKEKPFMKRQVQINDK-----DISEYKHAFFELI 517
 QY 526 LAYAGEVTVSYVGLDKNN-----DLFRNLKETMCSSMNPIMAOCPKSELSDKRPE 579
 Db 518 LKDGNSVIFSAKSAEKKNNMMWMAALISLOYSTLERM-----LDVTVLOEKEK- 564
 QY 580 TVATQPKMSLLOVELLNSKEPAYTRCIKPNDAKQGRFDE-----VLIRHOVXYLGL 632
 Db 565 -----EOMRILPSAEVYRFAEP-----DSEENILFEENWQKAGIPITIKAGTVLKL 609
 QY 633 MENLVRBRAGFAYRRKYEAFLORYKSLCPETWPMWAGRPDQGAVALVRHLGY--KPEEYKM 691
 Db 610 IERLTYHM--YADPNFVRFTLYTSFC-----RQELLSLILIEFIEPEPEPTEA 658
 QY 692 GRTKILFIRPKTLFATED-----SLEVR-----QSLATKIQAAMRGFHM--ROKFLRYK 739
 Db 659 DR-----IAIENDQPLSALKRFRREYIOPVOLRVLNCR--HWVEHHFYDFE 705
 QY 740 RSLICQSW--WRGTLGRKKAKRKAQAOTIRRLR-----GFIIRHSPRCPE 785
 Db 706 RQDIDLQRMEEFTIGY--RGKAMKKH--VSTITKIIOKKIARDNGRGHNTFQSSP--PT 760
 QY 786 NAEFLD---HYRASFLLNLR-----ROL-----PRNVLDTSWP-----TPPP 819
 Db 761 VEHMISRPQHIEFFDLTLHPLEIANQULLESPLYRAVOPELVSQVTKEDKEINSPN 820
 QY 820 ALR-----EASELLRELCMKMVMKYCRSI-- 844
 Db 821 LKMIHRTTNLTLMPKCIETENTLEERVAVVSRILEILOVFOELNNGVLEVVSAAMS 880
 QY 845 SPEWK-----OOL--OQKAVASEIFPKKKNY-----POSVP--RLFISTRIGT 884
 Db 881 SPVYRLDHTFEQPSROKKILEBAHLSHDHKKYLAIRLSINPCVPPFFGIYLTNILKT 940
 QY 885 EETISPVLOSIGSEPIQYA---VPVYKYDRKGYPKRPQOLLITPSAAVVIYEDAKYKQRI 940
 Db 941 EECNPEVYLRHKGELINFSKRRRVALETGEIOYQNOQPCYCLRVEPD-----IKRPF 991
 QY 941 DYANLTGISV--SSLSDSLVLVHQRDNKOK 970

Db 992 ENLNPWGNMKEKEFTDYLENKSLIEPRHDK 1022

Search completed: July 7, 2003, 14:25:39
 Job time : 43.4566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:21:30 ; Search time 43.3358 Seconds

(without alignments)
2770.284 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MRYRASALGSDGVRYTMESA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCT08_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1285.5	23.8	2058	9	US-09-815-379-17
2	1261.5	23.4	1016	12	US-10-044-303-2
3	1241.5	23.0	2048	9	US-09-815-379-12
4	1241.5	23.0	2057	9	US-09-815-379-10
5	1240	22.1	765	12	US-10-044-303-1
6	1192	22.1	1938	9	US-10-171-311-164
7	1192	22.1	1972	9	US-10-171-311-162
8	1188.5	22.0	1945	9	US-09-927-597-2
9	1188.5	22.0	1979	9	US-09-927-597-4
10	1186	22.0	569	10	US-09-925-300-1583
11	1175.5	21.8	845	9	US-09-927-597-14
12	1172	21.7	849	9	US-09-927-597-10
13	1172	21.7	852	9	US-09-927-597-12
14	1168.5	21.6	544	9	US-10-106-698-5864
15	1150.5	21.3	1285	9	US-10-205-823-273
16	1132	21.0	2099	9	US-10-128-714-3290
17	1130	20.9	2405	9	US-10-128-714-8290
18	1127	20.9	772	9	US-09-927-597-8
19	1105	20.5	2548	10	US-09-851-682A-1

20	1042	19.3	699	9	US-09-927-597-6	Sequence 6, Appl1
21	568	10.5	245	10	US-09-925-302-686	Sequence 686, App
22	566	10.5	570	10	US-09-764-853-551	Sequence 551, App
23	497	9.2	2568	10	US-09-866-108-3	Sequence 3, Appl1
24	244.5	4.5	294	9	US-10-106-698-5501	Sequence 5501, Ap
25	215.5	4.0	286	10	US-09-925-301-923	Sequence 923, App
26	207.5	3.8	237	9	US-09-925-299-937	Sequence 937, App
27	207.5	3.8	2549	9	US-09-950-634-3	Sequence 937, App
28	160	3.0	103	10	US-09-864-761-41932	Sequence 3, Appl1
29	150.5	2.8	103	10	US-09-864-761-41932	Sequence 34208, A
30	148	2.7	77	10	US-09-864-761-48341	Sequence 41932, A
31	145.5	2.7	1337	10	US-09-864-761-41891	Sequence 48341, A
32	144.5	2.7	100	10	US-09-864-761-39854	Sequence 1, Appl1
33	144.5	2.7	102	10	US-09-864-761-41762	Sequence 39854, A
34	144.5	2.7	134	9	US-10-091-438-227	Sequence 41762, A
35	144.5	2.7	134	10	US-09-764-853-760	Sequence 227, App
36	140	2.6	89	10	US-09-864-761-41891	Sequence 760, App
37	137.5	2.5	103	10	US-09-864-761-42604	Sequence 41891, A
38	136.5	2.5	103	10	US-09-864-761-42606	Sequence 42604, A
39	133.5	2.5	1045	10	US-09-815-242-10617	Sequence 42606, A
40	132.5	2.5	3899	9	US-10-171-311-4	Sequence 10617, A
41	132.5	2.5	3907	9	US-10-171-311-2	Sequence 4, Appl1
42	132.5	2.5	3917	9	US-10-171-311-8	Sequence 2, Appl1
43	132.5	2.5	3925	9	US-10-171-311-6	Sequence 8, Appl1
44	130.5	2.4	66	10	US-09-864-761-42111	Sequence 6, Appl1
45	130.5	2.4	66	10	US-09-864-761-42111	Sequence 42111, A

ALIGNMENTS

RESULT 1
US-09-815-379-17
Sequence 17, Application US/09815379
Publication No. US20030073613A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
APPLICANT: GERITSEN, MARY
TITLE OF INVENTION: ANGIOGENESIS ASSOCIATED PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE: 10716/35
CURRENT APPLICATION NUMBER: US/09/815, 379
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 60/191,134
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 17
LENGTH: 2058
TYPE: PRT
ORGANISM: Homo sapiens
US-09-815-379-17

Query Match 23.8%; Score 1285.5; DB 9; Length 2058;
Best Local Similarity 36.9%; Pred. No. 5.6e-90;
Matches 301; Conservative 150; Mismatches 250; Indels 115; Gaps 20;
QY 3 YRASALGSDGVRYTMESALTRADKGVQDFVLENTFSEANFLENLRRENLIYYIG 62
DB 44 YKOSTTHOKV-----TAMHPTNEEGVDMASTLELHG-GSIMYVNLFORRYKNOIYYIG 97
QY 63 PVLVSVPYRDLC-IYSROHMERRYRGVSFEVPHPLFAVADVYVRLTRERDQAVWSIG 121
DB 98 SILASVNPYQPIALYEPATMEQYSRHLGELPRPHIRAINECRCMKRYDQCIILSIG 157
QY 122 ESGGKTEAKTRILLQFAF-----TPADRGCAVDRLLQSNPVLAEFGNAKTYLRND 175
DB 158 ESGGKTESKRLIKFLSVISQSLLESLKESCVRAILLESSPIEAFGNKATYVNNN 217
QY 176 SSRGKYMDVQDFDEKGAPEVGHILSYLLEKSRVYHONHGERNHFVQLLEGDEEFLTR 235
DB 218 SSRGKYVQNLICQKNGIQCGRIVYDLLEKNRYVRNPGERNHYIFALLAGLEHEREE 277

```

OY 236 LGLENNPOSYLTVKGGCAKVSINDKSDKVMKRALSVIDFTEDEVEDLLSTIYASVTLH 295
DB 278 FYLS--TPENHYHLNOSGVEDKTIISDQSFREYITAMVMOFSKREVEVSRLLAGLHL 336
OY 296 GNIHFAEDSDNAQVTEENOLKYLRLRLEGSTLREALTFRKTIKAGEELLSPINLEOA 355
DB 337 GNIETIT--AGGAOVSEFKTALGRSHELLDPTQTLDTALTRSMFLREBELITPLNVOQA 394
OY 356 AAYARDALAKAVYSRTFTWLVRKIN--RSLASKDAESPMSRSTTVGLDITGEYFQHNSE 414
DB 395 VDSRSLAMALYACCFEWEWIKKINSRIGNEDFKS-----IGILIDFGEENEFVHF 446
OY 415 EPOFCINCEKIOOLFIELTLKSEQEYEAGEIAPEYOYFNKTIICDLVEEKFGLISI 474
DB 447 EOPNINYANEKLOEFYFNKHIISLEQLEYSREGUVEDIDMDNGECDDILIEKL--GLLAL 505
OY 475 LDEECLRPEATDITFLEKLEDTYVPHPHFLTKLADQTKRSILDRGEFRLLHYAGEVY 534
DB 506 IMESHFP--QATDSTLEKLIH--SQHANNHFV-----KPRVAVN--NEGVMHYAGEVOY 555
OY 535 SYTGFLDKNNDLFRUKETKMGSSANPIMAOE-----DKSELDDKKRPETVATOF 585
DB 556 DYRGILEKKNRDTFRDLMLLRESRDFIYDLFEHVSSNNODTLKCSKHRRPTVSSOF 615
OY 586 KMSILQVLEILRSKEPAYIRCIKPDNAKOPGRFEVYLIRHOVKYGLMELNLRVRAGPAY 645
DB 616 KSLHSLAMATLSSSNPFYRCIKPNQKMPDQDAVVLNOLARYGMLETVAIRAGYAV 675
OY 646 RRKYEAFLQRYKSLCPEYPMWAGRPQDGAVALVRHLGKPE-----687
DB 676 RRFODFYKRYK-----VLMRMLAL--PDVVRGKCTSLLOLYDASNS 715
OY 688 EYKMGRTKFIREFPKTLFATEDSLEYRQS-----LATK-----721
DB 716 EMQLGKTKVFLR-----ESLEOKLEKREKEEYSHAAMVTRAHVGLFLAKQYRKLYCYV 770
OY 722 -IOAANRGFMROKFLVRKSAICIOQSWRGTLGR 756
DB 771 IIOKNRFAFLRRRFLHLKKAIVFQKQJGQIARR 806

RESULT 2
US-10-044-303-2
; Sequence 2, Application US/10044303
; Patent No. US20020137161A1
; GENERAL INFORMATION:
; APPLICANT: Max-Planck-Gesellschaft e.V.
; TITLE OF INVENTION: Protein expression and structure solution using
; FILE REFERENCE: ST010209-EPA
; CURRENT APPLICATION NUMBER: US/10/044,303
; CURRENT FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1016
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: whole sequence
; OTHER INFORMATION: of recombinant protein M761-2R R238 E
; OTHER INFORMATION: The protein comprises as component (1) the aa
; OTHER INFORMATION: sequence of myosin II motor domain of
; OTHER INFORMATION: Dictyostelium, a three aa linker region and the
; OTHER INFORMATION: a-actinin aa sequence
US-10-044-303-2

Query Match 23.48; Score 1261.5; DB 12; Length 1016;
Best Local Similarity 34.38; Pred. No. 1.3e-88;
Matches 297; Conservative 181; Mismatches 322; Indels 67; Gaps 19;

OY 28 GVQDFVLENTSAAFTENLRRENRRLTYTIGPLVSVNRYRDLQIYSRQHMERYRG 87
||:| |::| | ||| |::| ||| |::| |::| |::| |::| |::| |::| |::| |::|

```

```

DB 91 GVEDMSEL--SYLNEPAVFNHRLRVYRNOPLLITYISGLFLVAVNPFKRIPIYQEMWDFKG 149
OY 88 VSEFVYPPHFAVDYRVALFRERRDQAVMISGESGAGTEATKRLQFAETCPAPR 147
DB 150 RRRENAVPHFALSDAVYKSMDDKQOSLITGESGAGTEATKRLQFAETCPAPR 209
OY 148 GGA--VBDRLIOSNPVLEAFNAKTYLNDNSSRFKAYMDVOFDEKCAPVGGHILSTLKE 205
DB 210 NGSGLVLOQLQANPILAEFGNAKTEFNNSRFKFIETIOFNSAGFISGASIOSYLKE 269
OY 206 SRVYHNGERNRHFVQQLLEGGEETRLRLGERNPQSLYLVKQCAKVSINDKSDM 265
DB 270 SRVFOSETENHIFVOLLGATAEKKLHL--AGPESNYLNOGSCVDIKVSSSEF 328
OY 266 KVMKALSVIDFTEDEVEDLLSTIYASVTLHGNTHFAEDSDNAQVTEENOLKYLRLIAY 325
DB 329 KITQADNDIVFQSEDEMSIFKTIAGILHGNIKFEKGAGEGAVLKDRTALNAASTVFGV 388
OY 326 EGTTLREALTHRKTIKAGEELLSPINLEOAAYARDALAKAVYSTFTWLVRKINRSLASK 385
DB 389 NPVLEKALMPEIRILAGRDVAOHLNVEKSSSHDALVKALYGLFLMVLKTIINVLCOE 448
OY 386 DAESPMSRSTVGLDITGEYFQHNSEPOFCINCEKIOOLFIELTKSEQEEYAE 445
DB 449 -----RKAVFTGVLDISGFETKVSFQDLCTINTEKLOQFPNNHMKLOEYLAKE 501
OY 446 GIANEPOY--FNKTIICDLVE--EKFGLISITDEECLRPEATDITFLEKLEDTYVPHPH 503
DB 502 KINMTFIDFGSDQATIDIGRQPGILALDQSVFP--NATDNTLITFL-----HSH 554
OY 504 FLTHKLADQTKRSILDRGEFRLLHYAGEVYSVYGFGLDKNNDLFRUKETKMGSSANPIM 563
DB 555 FSKKNAYKEEPRFS--KTEFGVTHYAGQVMTEIDMWLEKNDPLQOULELCFKDSSPNV 612
OY 554 AOCPEKSELSPDKRPE-----TVATQFKMSILQVLEILRSKEPAYIRCIKPDNAKOPGRFD 619
DB 613 TKLENDPNIASRAKKAFAETIVAQYKQOLASLWATLETTPHVRICILPNNKQLPKLE 672
OY 620 EYLIRHOVKYGLMELNLRVRAGPAYIRKRYAPALQRYKSLCPEYPMWAGRPQDGAVALV 679
DB 673 DKVVLQDLRCNGVLEGIRITRKGGPNRIYADFPVKRYLLAPNV--PDAEDSQATDAVL 731
OY 680 RHLGKPEEYKMGRTKFIREFPKTLFATEDSLEYRQSIAQANRGFMROKFLVRK 739
DB 732 KHLNDPQYRGTIKITFRAGQ--LARLEBARQRLGSEQKSD-----YLK 777
OY 740 RSLAICIOSMW-----RGLGRKKAARKWAOTIRRLIRGLFIRHSPPROPENAFLDHYRA 795
DB 778 RANELVQ--WINDKQASLESRDGDSISVQSFMAHREYKRTKPPRGQEVSELEAIYN 835
OY 796 SFLLNLR-----KQLRNVLDTSWPPPPALRPSLLR--ELCMKN-----NYWKY 840
DB 836 SLQTKRLRIKREPFAVAPAGLIPNETIDTSWALSLEKBAEOHABALRIELKROKKIAYLQKY 895
OY 841 CRSISPEWKOOLQOKAVASEIFKGRKD 867
DB 896 NRIL-----KLENNATTKSVYLSGNE 917

RESULT 3
US-09-815-379-12
; Sequence 12, Application US/09815379
; Publication No. US20030073613A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: GERRITSEN, MARY
; TITLE OF INVENTION: ANGIOGENESIS ASSOCIATED PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: 10716/35
; CURRENT APPLICATION NUMBER: US/09/815,379
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/191,134
; PRIOR FILING DATE: 2000-03-22

```

; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 2048
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-379-12

Query Match 23.08; Score 1241.5; DB 9; Length 2048;
 Best Local Similarity 36.1%; Pred. No. 1,4e-86;
 Matches 295; Conservative 154; Mismatches 252; Indels 117; Gaps 22;

```

OY 3 YRASALGSDGVRYTMSALTAARDVGVDFVLLNFTSQAFTENLRARRRENLITYIG 62
DB 41 YKOSTITHOKV-----TAMHPTBEGVDDMASTELHG-GSIMYNLFQKRNQIMTYIG 94
OY 63 PVLVSVPYRDLO-TYRQHMERYGVSEYEPHLEFAVDVYRALRTERDDAVMTSG 121
DB 95 SILASVNPYQPIAGLVEPATMEQYSRHLGELPRPHIFALANCYCRLMRHNOICILKG 154
OY 122 ESGAGTEATKRLLOFYAE-----TCPAERGAARDLLOSNPYLEAFNAKTLRDN 175
DB 155 ESGAGTEATKRLKFLSVISQSLSTLKECTSCVERAILLESSPIMEAFNAKTYNNN 214
OY 176 SSRGKYMVDQDFKCAPVGGHILSTYL-EKSRVYHONHGERNFHFVOLLGGEETLR 234
DB 215 SSRGKRVOLNIOGKQIGRIYDCILSSQNRVYRQNGERNYHIFALLAGLHEERE 274
OY 235 RLGIERNPOSYLYLVGQCAKVSINDSKDMYMRKALSVIDFTEDEVEDLLSYASVLH 294
DB 275 EFLYS-TPBNYHNLNOSGVEDKTIISDQSFREVITAMDVQFSKEEVEYSRLLAGILH 333
OY 295 LGNIHFAEDSDSAOYTTEQNLKYLTRLGVEGTLRLALTRKILKAGEELLSPLENG 354
DB 334 LGNIIEFT--AGCAQVSFTALGRSAELGLDPTQLTALQSRMFLGEEILTPLVNQ 391
OY 355 AAVARDALAKAVYSRTFTWLVRKIN-RSLASKADSPSRSTTVGLLDYGFVEYQHS 413
DB 392 AVDSRSLAMALYACCFEWNKINSRIKGNEDFS-----IGLIDIGCFENFEVNH 443
OY 414 FEOFCINVCNEKLOQLFIELTLKSEQEEYEAEGIAMPEVQYFNKKIICDLVEERKGIIS 473
DB 444 FEOFNINYANERKLOEYFNKHFSLQLEYSRGLWEDIDWIDNECIDLLEKTL-GLLA 502
OY 474 ILDECLRGCEATDLTFLEKLEDYKPRPHFLTHKLADQKTRKSLDRGEFRLHYAGEVT 533
DB 503 LINESHHP-QATDSTLEKILH-SQHANNHYV-----KPRVAVN--NFCVKHYAGEVQ 552
OY 534 YSVTGFELKNNDLFRNLKETWCSMNPIMAOCE-----DKSELSDKKRPETVATQ 584
DB 553 YVGRGILEKNRDTFDDLLNLRSERFIDYLFHVSRRNODTLCKGSHRRPTVSSQ 612
OY 585 FKM-SILOLVELILRSKEPAYIRCIKRPNDAKOPGREDEVILRHQVYGLIMENLRVAGF 643
DB 613 FKVDSLHSLMATLSSNPFYRCIKRPNQKMPDDQAVVNLQURSGMLETYVIRKAGY 672
OY 644 AYRKRYEAFLOKYSLCEPTWPMAGRPQDQAVLVRLGKPE-----687
DB 673 AYRRPFQDFYKRYK-----VLMRNLAL-PEDVAGKCTSLQLYDAS 712
OY 688 --EYKMGRTKIFIRPKTLFATEDSLEVRROS-----LATK-----721
DB 713 NSEMOUGTKYVFLR-----ESLEQLEKRRREEVSHAMVYIRAHVIGFLARKOYRVLYC 767
OY 722 ---IOAMRGFMHQKFLVYKRSALCLOSMMRGTLGR 756
DB 768 VVIIOIKRAFLRRRLHLKKAIVFOKOLRGQIARR 805
  
```

RESULT 4
 US-09-815-379-10
 ; Sequence 10, Application US/09815379
 ; Publication No. US20030073613A1

; GENERAL INFORMATION:
 ; APPLICANT: RASTELLI, LUCA
 ; APPLICANT: GERRITSEN, MARY
 ; TITLE OF INVENTION: ANGIOGENESIS ASSOCIATED PROTEINS AND NUCLEIC ACIDS
 ; TITLE OF INVENTION: ENCODING THE SAME
 ; FILE REFERENCE: 10716/35
 ; CURRENT APPLICATION NUMBER: US/09/815,379
 ; CURRENT FILING DATE: 2001-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,134
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 2057
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-379-10

Query Match 23.08; Score 1241.5; DB 9; Length 2057;
 Best Local Similarity 36.1%; Pred. No. 1,4e-86;
 Matches 295; Conservative 154; Mismatches 252; Indels 117; Gaps 22;

```

OY 3 YRASALGSDGVRYTMSALTAARDVGVDFVLLNFTSQAFTENLRARRRENLITYIG 62
DB 41 YKOSTITHOKV-----TAMHPTBEGVDDMASTELHG-GSIMYNLFQKRNQIMTYIG 94
OY 63 PVLVSVPYRDLO-TYRQHMERYGVSEYEPHLEFAVDVYRALRTERDDAVMTSG 121
DB 95 SILASVNPYQPIAGLVEPATMEQYSRHLGELPRPHIFALANCYCRLMRHNOICILKG 154
OY 122 ESGAGTEATKRLLOFYAE-----TCPAERGAARDLLOSNPYLEAFNAKTLRDN 175
DB 155 ESGAGTEATKRLKFLSVISQSLSTLKECTSCVERAILLESSPIMEAFNAKTYNNN 214
OY 176 SSRGKYMVDQDFKCAPVGGHILSTYL-EKSRVYHONHGERNFHFVOLLGGEETLR 234
DB 215 SSRGKRVOLNIOGKQIGRIYDCILSSQNRVYRQNGERNYHIFALLAGLHEERE 274
OY 235 RLGIERNPOSYLYLVGQCAKVSINDSKDMYMRKALSVIDFTEDEVEDLLSYASVLH 294
DB 275 EFLYS-TPBNYHNLNOSGVEDKTIISDQSFREVITAMDVQFSKEEVEYSRLLAGILH 333
OY 295 LGNIHFAEDSDSAOYTTEQNLKYLTRLGVEGTLRLALTRKILKAGEELLSPLENG 354
DB 334 LGNIIEFT--AGCAQVSFTALGRSAELGLDPTQLTALQSRMFLGEEILTPLVNQ 391
OY 355 AAVARDALAKAVYSRTFTWLVRKIN-RSLASKADSPSRSTTVGLLDYGFVEYQHS 413
DB 392 AVDSRSLAMALYACCFEWNKINSRIKGNEDFS-----IGLIDIGCFENFEVNH 443
OY 414 FEOFCINVCNEKLOQLFIELTLKSEQEEYEAEGIAMPEVQYFNKKIICDLVEERKGIIS 473
DB 444 FEOFNINYANERKLOEYFNKHFSLQLEYSRGLWEDIDWIDNECIDLLEKTL-GLLA 502
OY 474 ILDECLRGCEATDLTFLEKLEDYKPRPHFLTHKLADQKTRKSLDRGEFRLHYAGEVT 533
DB 503 LINESHHP-QATDSTLEKILH-SQHANNHYV-----KPRVAVN--NFCVKHYAGEVQ 552
OY 534 YSVTGFELKNNDLFRNLKETWCSMNPIMAOCE-----DKSELSDKKRPETVATQ 584
DB 553 YVGRGILEKNRDTFDDLLNLRSERFIDYLFHVSRRNODTLCKGSHRRPTVSSQ 612
OY 585 FKM-SILOLVELILRSKEPAYIRCIKRPNDAKOPGREDEVILRHQVYGLIMENLRVAGF 643
DB 613 FKVDSLHSLMATLSSNPFYRCIKRPNQKMPDDQAVVNLQURSGMLETYVIRKAGY 672
OY 644 AYRKRYEAFLOKYSLCEPTWPMAGRPQDQAVLVRLGKPE-----687
DB 673 AYRRPFQDFYKRYK-----VLMRNLAL-PEDVAGKCTSLQLYDAS 712
OY 688 --EYKMGRTKIFIRPKTLFATEDSLEVRROS-----LATK-----721
DB 713 NSEMOUGTKYVFLR-----ESLEQLEKRRREEVSHAMVYIRAHVIGFLARKOYRVLYC 767
  
```

OY 722 ---10AAMRGPHMROKELRYKRSALCISQMSWGGTLGR 756
Db 768 VVIOKMYRAFLKRRRLHLKKAIVFQKOLRGQIARR 805

RESULT 5
US-10-044-303-1

; Sequence 1, Application US/10044303
; Patent No. US20020137161A1
; GENERAL INFORMATION:
; APPLICANT: Max-Planck-Gesellschaft e.V.
; TITLE OF INVENTION: Protein expression and structure solution using
; TITLE OF INVENTION: Specific fusion vectors
; FILE REFERENCE: ST010209-EPA
; CURRENT APPLICATION NUMBER: US/10/044,303
; CURRENT FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Partial myosin
; OTHER INFORMATION: sequence of Dictyostelium: Component (1) of the
; OTHER INFORMATION: recombinant protein M761-2R R238E
US-10-044-303-1

Query Match 23.0%; Score 1240; DB 12; Length 765;
Best Local Similarity 38.1%; Pred. No. 4e-87;
Matches 265; Conservative 154; Mismatches 248; Indels 28; Gaps 12;

OY 28 GVQDFVLENTSEAFATENTLRRENNLYTYIGPVLSVNPYRDLOIYSROHMERIRG 87
Db 91 GVEDMSEL-STYNEPAVPHNLRYRQNDLIYSGFLVAVNPKRIPRIYTOEWVDFKG 149
OY 88 VEFYEPHPLFAVADTVYRALTERRDQAVMISGSGAGTEAKRRLQFAEPCAPER 147
Db 150 RRNEVAPHFIAISVAPRSMDDKONOSLITGSGAGTEAKRRLQFAEPCAPER 209
OY 148 GGA--VRDLQSNVLEAFNAKTLRNDNSRFQKYMDOVDFKAPGVGHILSYLLEK 205
Db 210 NSGVLEQOILQANPLLEAFNAKTLRNNNSRFQKYEIOFNSAGFISGASIOSYLLEK 269
OY 206 SRVYHONGERNFHYFQLLLEGSEETLRGLERNPOSYLYVKGCAKYSINDKSDW 265
Db 270 SRVYFQSETERNYHIFQYLLAGATAEKKALHL-AGPESFNLNOSGCVDIKGYSDSEF 328
OY 266 KVMRKALSYIDFTEDEVEDLLSIVASVLIHGINHFADEDSNAQVTTENOLKYLRLLG 325
Db 329 KIRQAMDIVGSGEOMKIFKILGILHGNKREKAGGAVLAKDTALNAASYGV 388
OY 326 EGTTLREALTHRKIAKGEELISPLNLEQAAVADALAKAYSRFTVLVKIRSLASK 385
Db 389 NPVELKALMEPRILLAGRDVAQHINVERSSSRDALVKAIFYGLFVLMVKINNVCOE 448
OY 386 DASPMSRSTYVGLDIYGFVFOHNSFQCIYCNKELQOLFTELKSEOEYEA 445
Db 449 -----RAYFEGVLDISGFELFKVNSFQCIYCNKELQOLFTELKSEOEYEA 501
OY 446 GAAEVPQY-FNNKILICDVE-EKFGKISILDECLRPGEATDLPLEKLEDTYVKKPH 503
Db 502 KIMWTFIDGLDSQATIDLDGQPGIALLDEQSVFP-NATDWTLITKL-----HSH 554
OY 504 FLTHKLAOKTRKSLDRGEFRLIHYAGEVTVSGFELDKNNDLERNLKEMCSMNPI 563
Db 555 FSKRNKAYEPRPS--KTEBQVHYAGVYVEIQMLEKNKPLDQDDLELCKSSDVA 612
OY 564 AOCFDSKSLDKKRP-----TVATQFKMSLQIVELIRSKEPAYIRKICIKPNDAKOGRFD 619
Db 613 TKLPNDPNIASRAKKGANFTYVAOYKQELASIMATLETITNHFVRCIIPNNKQLPATLE 672

OY 620 EVLIRHVKYKILGIMENLRVRAAGFAYRKYEAFLQKSLCPETWPMNAGRPQDQAVLV 679
Db 673 DKVYDQLRONGVLEGIRIRTKGFPNNKIYADFKYIILANV-PROADEQSKATDAVL 731
OY 680 RHLGKPEEYKMGRTKIFIRFPKTLFATEDSLLEVR 714
Db 732 KHLNIDPEQYRFGTITKIFRAGQ-LARIEARQR 765

RESULT 6
US-10-171-311-164

; Sequence 164, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kametkar, Shubhang
; APPLICANT: Glati, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersht, Sebastien
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-164

Query Match 22.1%; Score 1192; DB 9; Length 1938;
Best Local Similarity 29.5%; Pred. No. 8.7e-83;
Matches 316; Conservative 197; Mismatches 357; Indels 200; Gaps 30;

OY 40 SEAFATENTLRRENNLYTYIGPVLSVNPYRDLOIYSROHMERIRGVSFEYEPHPLFA 99
Db 97 NEASVLIHNRKRYFSGLIYTYGSLFCVVPYKHLPIYSEKIVMYGKRKHEMPHIA 156
OY 100 VADTVYRALTERRDQAVMISGSGAGTEAKRRLQFAEPCAPER-----GAVR 152
Db 157 IADTAARSMLODREDSITCTGSEAGAGTEAKRRLQFAEPCAPER-----GAVR 215
OY 153 DRLOSNPVLEAFNAKTLRNDNSRFQKYMDOVDFKAPGVGHILSYLLEKSVYHON 212
Db 216 KOLQANPLLEAFNAKTLRNDNSRFQKYEIRINDVGYIVGANIETYLLEKSRARIOA 275
OY 213 HGERNHVYQLLLEGSEETLRGLERNPOSYLYVKGCAKYSINDKDKVMKAL 272
Db 276 RDRERTHIFYIMAGAKEMRSDLLLE-GFNNTYPLSG-FVPIPAADDDMEFOEYVAM 333
OY 273 SVYDFTEDEVEDLLSIVASVLIHGINHFADEDSN-AQVTTENOLKYLRLLGEGTTL- 330
Db 334 AIMGSEERQSLIKVAVSSVLQGNIVRKENTQASMPONTAOKYCHLMGINVDET 393
OY 331 REALTHRKIAKGEELISPLNLEQAAVADALAKAYSRFTVLVKIRSLASKDAES 389
Db 394 RSLITLR--TKGRDVYQAQKTEQADFEVLAATYERLFRWILTFVNNALDKTRHOG 451
OY 390 PSNRSTTVGLDIYGFVFOHNSFQCIYCNKELQOLFTELKSEOEYEAEGIAW 449
Db 452 ASF-----IGIIDLIAGFELFVNSFQCIYCNKELQOLFTELKSEOEYEAEGIAW 506


```

QY 930 IVEDAKVORIDYANLTG-1SVSLSDSLFLVHVEDNKGKGDVYLOS DHYIETLTAKTA 988
Db 959 LEEDEAROKIOLKYEYTAEKIKKLDELIVMDQ--NNKLSKERKLEERSIDLTMLA 1016
QY 989 LSADRVNININOGSITFAGPGRGDIIDFTSGSELLITKANGHAYVA 1038
Db 1017 EEEERAKN-----LTKLNKKEHSMTS 1037

RESULT 8
US-09-927-597-2
; Sequence 2, Application US/09927597
; Publication No. US20030032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Betraud, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYTOPO18
; CURRENT APPLICATION NUMBER: US/09/927,597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1945
; TYPE: PRT
; ORGANISM: Human
US-09-927-597-2

Query Match 22.0%; Score 1188.5; DB 9; Length 1945;
Best Local Similarity 29.3%; Pred. No. 1.6e-82;
Matches 316; Conservative 197; Mismatches 357; Indels 207; Gaps 30;

QY 40 SEAFIENLRFRRENLIYTYIGPVLSVNPYRDLQIYSRQHMERYGVSFEVPPHLEA 99
Db 97 NEASVLAHLNREYFSGLIYTSGLFCVYVNPYKHLPIYSEKIDYMGKKRHHMPPIYA 156
QY 100 VADTVYRALRTERRDQAVMISGSGAGTEATKRLLOFYATCPAPERG----- 148
Db 157 IADTAYRSMLODRQOSITCTGSEAGAGTEATKRLLOFYATCPAPERG----- 148
QY 149 ---GAVRDLLOSNPVLEAFGNAKTLRDNSSRFKGYMDVOFDPKGAAPVGGHILSYLLEK 205
Db 216 FAYGELERKOLQANLLEAFGNAKTLRDNSSRFKGYMDVOFDPKGAAPVGGHILSYLLEK 205
QY 206 SRVHQNHGERNFHFYOLLESGEETLRRLGLEERNFOSYLYLVKGCAYSSINDKSDW 265
Db 276 SRAIRQANDERTFHFYIMAGAKKMRSDILLE-GFNNTYFLSNG-FVPIPAADDEMF 333
QY 266 KVMKALSYIDTDEVEDLISYASVYHLGNIHPADDEDEN-AQVITENQKLYTLRLG 324
Db 334 QETVEAMAIMGSEEBEOLSLKVSVYQLGNIVFKERNDDQMSPDNRAAQVCHIMG 393
QY 325 VEGTTL-REALTFRKLIAGBELSLPLN-EOAAVARDALAKAVYSRTFTVLVRKINSL 382
Db 394 INVTDTTRSIILPR--IKVGRDVOYKQTKEDQADRAVALAKATYERLFRWILTRVKAL 451
QY 383 ASKDAESPSWSTYVLGLIDYGFYFQHNSEFCPCINCKEKLQOLFIELTLKSDEEY 442
Db 452 DKTRHOGASF-----LGIIDTAGFEIPEVNSFEQCIANYTEKLOQLFNTHTMLEDEY 506
QY 443 EAGTMEPVOYFENKNTIC-DLVEEKFK--GIISLDEECLRPGEATDLTLEKLEDTVK 499
Db 507 QREGIEMWNTIDGLDQPCIELEIRPNPFGVALLDDECFMR-KATDKSVVEKLCSEOG 565
QY 500 PPHPLTFLHKLADQKTRSLDRGEFRLILAYAGEVTVSVTGLDKNDLLFRULKETSMSM 559
Db 566 SHPKF-----QKPKOLKDKTEFESIHYAGKVDYNASAWILTKNDPILNDVYTSILNASS 618

```

```

QY 560 NPIMACF-----DKSELSOKRP-----EYVAFQFKMSLLOLVELLRSK 599
Db 619 DKFVADLMKVDRIYGLDOMAKTESSLPSASAKTKGFRYVGOYLKQOLKIMLTTLT 678
QY 600 EPAYIRCIKPNDAKQGRDEVLIRHVOYGLIMENLVRRAAGVYRRKRYEAFLOKRYSL 659
Db 679 TPNFVRILIPNHEKRSGLDAFLVYQLQNCVLEIGIRICRQGFQFNRRVYFQEFQREIL 738
QY 660 CPETWPMWAGRPDQAVAVLVRHLGYKPEEYKGRTRKIFRPFKTLFATEDSILEVROSIA 719
Db 739 AANAIPKGFMDGKQACILMTKALELDPNIRIGQSKIFPR-TGVLAHLEERDLKITYYI 797
QY 720 TKIOAWRGFMRKQKFLVYKRSALICIO-----SWMR----- 750
Db 798 MAFQAMCGYLAIRKAKFAKROQOLAKVIORNCAYATKLRMQWNRALFTKYVLOVYRQ 857
QY 751 -----GTJGRKKAARKKMAAOTIRLRIGLIRHSPRCENNAFFLDHVASFLMLR 802
Db 858 EEMQAKEDLOKTEKEROOK-AENELKELEQ--KHSOLTEE----- 895
QY 803 RQLPNRVLDTSWPPPPALREASELLRELCKMMVKYCRSISPEWKOLOQKAVASEIF 862
Db 896 ---KNLLOEOLDAETLYAEAEEMVRLAAK-----KQELE--LIHE-- 933
QY 863 KKKNDYPOQSVPLRFLISTRLGTEISPRVLOSISEPIOTAVPVYKDKKGRPRROL 922
Db 934 -----MEARLEEBEDRQOIOA-----EKK-KMAQOMLD 961
QY 923 LTPSAVITYEDAKVORIDYANLTG-1SVSLSDSLFLVHVEDNKGKGDVYLOS DHYI 981
Db 962 LEEQ--LEEEAROKIOLKYEYTAEKIKKLDELIVMDQ--NNKLSKERKLEERSI 1016
QY 982 ETLTKTALSADRVNININOGSITFAGPGRGDIIDFTSGSELLITKANGHAYVA 1038
Db 1017 DLTFTLAEERAKN-----LTKLNKKEHSMTS 1044

RESULT 9
US-09-927-597-4
; Sequence 4, Application US/09927597
; Publication No. US20030032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Betraud, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYTOPO18
; CURRENT APPLICATION NUMBER: US/09/927,597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1979
; TYPE: PRT
; ORGANISM: Human
US-09-927-597-4

Query Match 22.0%; Score 1188.5; DB 9; Length 1979;
Best Local Similarity 29.3%; Pred. No. 1.7e-82;
Matches 316; Conservative 197; Mismatches 357; Indels 207; Gaps 30;

QY 40 SEAFIENLRFRRENLIYTYIGPVLSVNPYRDLQIYSRQHMERYGVSFEVPPHLEA 99
Db 97 NEASVLAHLNREYFSGLIYTSGLFCVYVNPYKHLPIYSEKIDYMGKKRHHMPPIYA 156
QY 100 VADTVYRALRTERRDQAVMISGSGAGTEATKRLLOFYATCPAPERG----- 148
Db 157 IADTAYRSMLODRQOSITCTGSEAGAGTEATKRLLOFYATCPAPERG----- 148
QY 149 ---GAVRDLLOSNPVLEAFGNAKTLRDNSSRFKGYMDVOFDPKGAAPVGGHILSYLLEK 205

```



```

Db      507 QREGIENMFIDGLDPCIELIERPNNPCGVALLDDECEWFP-KATDKSFVEKCTEG 565
Qy      500 PHPEFLTHKLADQTRKSLDGRGRLHAGEVYVSTGFLDKNDLLFRNLKETMCSM 559
Db      566 SHRF-----QKPKOLDKTEFSIIHYAGKVDYNASWLTNNMDPLNDVNTSLNMS 618
Qy      560 NPMAOCF-----DKSELSDKRRP-----ETVATOFKMSILOVELIRSK 599
Db      619 DKFVADLMKDVRIVGLDQAKMTESSLPSSAKTKKGMRTYGOULYKGLKMTTLRNT 678
Qy      600 EPATRICIKPNDKQOPREDEVILRHQVYLGMLNLRVRAGFAVRRKYEAFLQRYSL 659
Db      679 TPWFVRCIIPNHEKRSGLDAFLVLEOLRCNGVLEGIRICROGFPNRIYFOEFRQREYL 738
Qy      660 CETWPMWAGRPDQGVAVLVRLHGYKPEYKMGRTKFIREFPKTLFATEDSLVEKROSLA 719
Db      739 AANAIPKGFMDGKACILMTKALEDDPNLYRIGOSKIFFR-TGVLAHLEERDLKITDVI 797
Qy      720 TKIOAMRGFHWKQFLVRKRSALCIO-----SMWR 750
Db      798 MAFOAMCRGYLARKAFARQOOLTMKVIORNCAYLKLNNQWNR 843

```

```

RESULT 13
US-09-927-597-12
/ Sequence 12, Application US/09927597
/ Publication No. US20030032018A1
/ GENERAL INFORMATION:
/ APPLICANT: Malik, Fady
/ APPLICANT: Beraud, Christophe
/ APPLICANT: Freedman, Richard
/ APPLICANT: Craven, Andrew
/ APPLICANT: Sakowicz, Roman
/ APPLICANT: Hartman, James
/ TITLE OF INVENTION: Human smooth muscle myosin heavy chain
/ FILE REFERENCE: CYTO018
/ CURRENT APPLICATION NUMBER: US/09/927,597
/ CURRENT FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 852
/ TYPE: PRT
/ ORGANISM: Human
US-09-927-597-12

```

```

Query Match      21.7% Score 1172; DB 9; Length 852;
Best Local Similarity 35.0% Pred. No. 8.5e-82;
Matches 268; Conservative 157; Mismatches 267; Indels 74; Gaps 17;

Qy      40 SEAAFIENLRRERENLIYIGPVLYSVNPRDLOIYSGRQHMERYRGVFEVPPHLEA 99
Db      97 NEASVYLNHREYFSLIYISGLPCVVVNPYKHLPIYSEKLYDMKKGKREMPHITA 156
Qy      100 VADTVYRALRTERROQAVMISGSGAGKTEATKRLQFYAETCPAPERG----- 148
Db      157 IADTAIRSMLODRDOSILCTGSGAGKTEATKRYOYLA-VVASSHKGKDKTSTIGPS 215
Qy      149 ---GAVRDLQSNPYLEAFGNNAKTLRNDSRFGYMYODDFGAPRGHIIISLEK 205
Db      216 FAYGELEKOLLQANPLEAFGNNAKTVKNDSSRFGEFIRINDVGYIVGAVIETYLEK 275
Qy      206 SRVYVHONHGRNFHFYQLLEGGEEETLRLGLERNPOSYLYLVGOCAKVSSINDKSDM 265
Db      276 SRAIRQARDETFHFYIYIMAGAKEMRSDILLE-GFNNTYPLSNG-FYPIPADDEMEF 333
Qy      266 KVMKRALSYIDFTEVEDELLSIVASVHLGNIHAADEDSN-AQVTTENOLKYLTRLIG 324
Db      334 QCTVEAMATMGFSEEDQSLIKLVSSVQLGIVFKKEKNDQASMPNTAAQKCHLMG 393
Qy      325 VEGTTL-REALTNRKTIAGKEELLPLNL-EQAAVARDALAAVYSRTFTMLVRKINSL 382

```

```

Db      394 INTDTRSLITFR--IKVGRDVQAKQTEQADFAVEALAAATYERLEFRWLITRVNKA 451
Qy      383 ASKDAESPMSRTTVGLLDIYEFGEFOHNSFEQFCINYNKRILOULFIELTKSPQREY 442
Db      452 DKTHRGASP-----LGIDIAFGELFEVNSPQDCINTNKTQOLFPHMTFIEQREY 506
Qy      443 EAGGIMEVEYQVYFNKNTIC-DLYEEKF--GIISILDECLAPGEATDLPLEKLEDPVK 499
Db      507 QREGIENMFIDGLDQPCIELIERPNNPCGVALLDDECEWFP-KATDKSFVEKCTEG 565
Qy      500 PHPEFLTHKLADQTRKSLDGRGRLHAGEVYVSTGFLDKNDLLFRNLKETMCSM 559
Db      566 SHRF-----QKPKOLDKTEFSIIHYAGKVDYNASWLTNNMDPLNDVNTSLNMS 618
Qy      560 NPMAOCF-----DKSELSDKRRP-----ETVATOFKMSILOVELIRSK 599
Db      619 DKFVADLMKDVRIVGLDQAKMTESSLPSSAKTKKGMRTYGOULYKGLKMTTLRNT 678
Qy      600 EPATRICIKPNDKQOPREDEVILRHQVYLGMLNLRVRAGFAVRRKYEAFLQRYSL 659
Db      679 TPWFVRCIIPNHEKRSGLDAFLVLEOLRCNGVLEGIRICROGFPNRIYFOEFRQREYL 738
Qy      660 CETWPMWAGRPDQGVAVLVRLHGYKPEYKMGRTKFIREFPKTLFATEDSLVEKROSLA 719
Db      739 AANAIPKGFMDGKACILMTKALEDDPNLYRIGOSKIFFR-TGVLAHLEERDLKITDVI 797
Qy      720 TKIOAMRGFHWKQFLVRKRSALCIO-----SMWR 750
Db      798 MAFOAMCRGYLARKAFARQOOLTMKVIORNCAYLKLNNQWNR 843

```

```

RESULT 14
US-10-106-698-5864
/ Sequence 5864, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypep
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ NUMBER OF SEQ ID NOS: 5864
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 5864
/ LENGTH: 544
/ TYPE: PRT
/ ORGANISM: Homo sapiens
FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (266)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (321)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (527)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (528)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (529)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5864

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:12:35 ; Search time 53.1849 Seconds

(without alignments)
1887.083 Million cell updates/sec

Title: US-09-893-371-1

Perfect score: 5398

Sequence: 1 MYRNASALGSDGVRTWESA.....LITKAKNGHLAVVAPRLNSR 1044

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR73:
2: PIR1:
3: PIR2:
4: PIR3:
5: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	5261	97.5	1028	2	S37146 myosin I heavy chain
2	5143	95.3	1028	2	A59253 myosin I beta - hu
3	5128	95.0	1028	2	S41749 myosin heavy chain
4	4371	81.0	1028	2	I51173 myosin I beta - bu
5	4167	77.2	807	2	H75634 myosin-IC - mouse
6	2632	48.8	1026	2	S45574 myosin IB - fruit
7	2252	41.7	448	2	B45438 myosin I beta, MMI
8	2109	39.1	1078	2	B45439 myosin-I, Myr 1c (
9	2106.5	39.0	1043	1	A29483 myosin-I, Myr 1b (
10	2094.5	38.8	1107	2	C45439 myosin I alpha cha
11	2081.5	38.6	1079	2	A45438 myosin I heavy chain
12	2080	38.5	1136	2	A45439 myosin heavy chain
13	2075.5	38.4	1094	2	S32404 myosin heavy chain
14	1904	35.3	1000	1	A33620 myosin heavy chain
15	1861	34.5	1006	1	A33933 myosin I myr 4 - r
16	1803.5	33.4	1011	2	S45573 myosin IA - Caenor
17	1756	32.5	1017	2	T24349 myosin IA - mouse
18	1739	32.2	1003	2	S33760 myosin IB heavy ch
19	1726.5	32.0	1099	1	A59300 myosin-IC (similar
20	1719.5	31.9	1099	1	S31926 myosin I heavy cha
21	1715	31.8	1109	2	S33601 myosin I heavy cha
22	1712.5	31.7	1107	1	S32517 myosin heavy chain
23	1692.5	31.4	1111	1	A33284 myosin heavy chain
24	1601	29.4	994	1	J00151 myosin heavy chain
25	1588.5	29.0	1147	1	MMAX1B myosin heavy chain
26	1567.5	28.6	1100	2	A56511 myosin I myoA - Em
27	1564	28.6	1249	2	A56511 myosin heavy chain
28	1545	28.6	1168	1	MMAX1C myosin heavy chain
29	1523.5	28.2	1168	1	MMAX1C myosin heavy chain

30	1509.5	28.0	1217	2	T39427	probable myosin I
31	1501	27.8	1219	2	S54570	probable membrane
32	1473	27.3	1215	2	T32734	myosin-IA - Acanth
33	1453	26.9	1181	2	T30578	myosin IC - slime
34	1450	26.9	1271	2	S37958	myosin heavy chain
35	1382.5	25.6	1446	2	T04528	myosin heavy chain
36	1361	25.2	2357	2	A59249	class VII unconvn
37	1354	25.1	1203	2	A59257	myosin VIIa, short
38	1354	25.1	2175	2	A59255	myosin VIIa, long
39	1352	25.0	1515	2	S51824	myosin heavy chain
40	1346	24.9	839	2	F83334	myosin-like protei
41	1336	24.7	1528	2	T14279	myosin heavy chain
42	1335	24.7	2121	2	A59233	myosin heavy chain
43	1331	24.7	1477	2	T00957	myosin VII-like pr
44	1331	24.7	2062	2	A59297	myosin heavy chain
45	1328	24.6	1611	2	A84743	myosin X - mouse
						probable myosin he

ALIGNMENTS

RESULT 1

S37146 myosin I heavy chain - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 02-Feb-2001

C:Accession: S37146

R:Ruppert, C.; Gogel, J.; Reinhard, J.; Baehler, M.

submitted to the EMBL Data Library, August 1993

A:Description: MYR-2, a novel class-I Myosin identified in rat brain.

A:Reference number: S37146

A:Accession: S37146

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1028 <RUP>

C:Superfamily: EMBL:X74800; NID:9400428; PION:CA52807.1; PID:9400429

C:Keywords: brush border myosin heavy chain I; myosin motor domain homology

F:14-683/Domain: myosin motor domain homology <MMOT>

F:105-112/Region: nucleotide-binding motif A (P-loop)

Query Match 97.5% Score 5261; DB 2; Length 1028;
Best Local Similarity 98.7% Pred. No. 0;
Matches 1015; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY	17	MESALTARDVGVODFVLENTSEAFTEMLRRFRRENLITYIGPVYVNPYRDLQI	76
DB	1	MESALTARDVGVODFVLENTSEAFTEMLRRFRRENLITYIGPVYVNPYRDLQI	60
QY	77	YSRQHERYRGVSFEVPPHFAVDYVYRALTERRRQAVMIGESGAGTETKRLQ	136
DB	61	YTRQHERYRGVSFEVPPHFAVDYVYRALTERRRQAVMIGESGAGTETKRLQ	130
QY	137	FPAETCPAPERGAVERDLQSNPYLEAFNAKTLRNDSSRFKQMDVDFGAPVGG	196
DB	121	FPAETCPAPERGAVERDLQSNPYLEAFNAKTLRNDSSRFKQMDVDFGAPVGG	180
QY	197	HLSTYLKSRVYVHNGERNFHPVYQLLEGGEETLRRLGLENNPOSYLYLVGQCAV	256
DB	181	HLSTYLKSRVYVHNGERNFHPVYQLLEGGEETLRRLGLENNPOSYLYLVGQCAV	240
QY	257	SSINDKSPMKVYRKALSYIDFTEDEVEDLTSIVASVHLGNIHFAAEDSNAOVTENOL	316
DB	241	SSINDKSPMKVYRKALSYIDFTEDEVEDLTSIVASVHLGNIHFAAEDSNAOVTENOL	300
QY	317	KLTFTLLVEGEGTLREALTHKRIIAKGEELISPLNEDQAAVARDALAKAVYSRTFTWLV	376
DB	301	KLTFTLLVEGEGTLREALTHKRIIAKGEELISPLNEDQAAVARDALAKAVYSRTFTWLV	360
QY	377	KINRSLSKADSPSMRSTVYGLLDIYGFEVFOHNSFEGFCINYNCEKIQOIFIELTLK	436
DB	361	KINRSLSKADSPSMRSTVYGLLDIYGFEVFOHNSFEGFCINYNCEKIQOIFIELTLK	420

```

OY 437 SEOEYEAEGIAEPVOYFNNKIIICDLVEEKFKGIIISILDEECLRGEATDLTFLEKLED 496
DB 421 SEOEYEAEGIAEPVOYFNNKIIICDLVEEKFKGIIISILDEECLRGEATDLTFLEKLED 480
OY 497 TVKPHPHFLTHKLADOKTRKSLDRGEFRLHYAGEVTVYATGFLDKNNDLFFNLKETMC 556
DB 481 TIKHHPHFLTHKLADOKTRKSLDRGEFRLHYAGEVTVYATGFLDKNNDLFFNLKETMC 540
OY 557 SSNMPMAOCPSKSELSDDKRPETVATQFKMSLLQVLEILRSKPEATYIRCIKPNDAKOPG 616
DB 541 SSTPMIAOCPSKSELSDDKRPETVATQFKMSLLQVLEILRSKPEATYIRCIKPNDAKOPG 600
OY 617 RFDEVLIRHQQVLYGLMENLVRBRAGFAYRRKYEAFLQRYKSLCPEWPMAGRPDQGYA 676
DB 601 RFDEVLIRHQQVLYGLMENLVRBRAGFAYRRKYEAFLQRYKSLCPEWPMAGRPDQGYA 660
OY 677 VLVHRLGKPEEYKMGRTKIFIRPKTLFATEDSELEVRROSLATKIOAANGCFHMRQKFL 736
DB 661 VLVHRLGKPEEYKMGRTKIFIRPKTLFATEDSELEVRROSLATKIOAANGCFHMRQKFL 720
OY 737 RVKRSALCISQMMWGTIGRRKAARKWAQAOTIRRLINGFLRHSRCPENAFELDHVRAS 796
DB 721 RVKRSALCISQMMWGTIGRRKAARKWAQAOTIRRLINGFLRHSRCPENAFELDHVRAS 780
OY 797 FLNLRLRPLPNNVLDTSWPTPPALREASELLRELCKMNMWYKCRSISPEMKOOLQOKA 856
DB 781 FLNLRLRPLPNNVLDTSWPTPPALREASELLRELCKMNMWYKCRSISPEMKOOLQOKA 840
OY 857 VASELFFKCKNDYPOSVPRFLFSTRLGTEESIPRYLOSLGSEPIQYAVPVYKDYRKGYK 916
DB 841 VASELFFKCKNDYPOSVPRFLFSTRLGTEESIPRYLOSLGSEPIQYAVPVYKDYRKGYK 900
OY 917 RPRQLLTPSAVVIYEDAKVKQRIDYANLGTISVSSLSDSLFLVHVOQEDNKKQGDVYLO 976
DB 901 RSRQLLTPSAVVIYEDAKVKQRIDYANLGTISVSSLSDSLFLVHVOQEDNKKQGDVYLO 960
OY 977 SDHVETLTKTALSADRYNNININGSTTFAGGPGRGDIIDFTSGSELLITKAKNGHLAV 1036
DB 961 SDHVETLTKTALSADRYNNININGSTTFAGGPGRGDIIDFTSGSELLITKAKNGHLAV 1020
OY 1037 VAPRLNSR 1044
DB 1021 VAPRLNSR 1028

RESULT 2
A:Species: Homo sapiens (man)
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
C:Accession: A59253
R:Cloned, F.: Amaraoui, A. E.; Blanchard, S.; Lenolir, M.; Ripoll, C.; Vago, P.; Hamel, C.;
Genomics 40, 332-341, 1997
A:Title: Cloning of the genes encoding two murine and human cochlear unconventional type
A:Reference number: A59253; MIMD:97237053; PMID:9119401
A:Accession: A59253
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 11028 <CRO>
A:Cross-references: GB:X98507; NID:g1926310; PIDN:CA67131.1; PID:g1926311
A:Experimental source: dev stage adult; tissue type kidney
C:Genetics:
A:Gene: myo-1b
A:Map position: 17p3.2-p13.3
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
F:14-683/Domain: myosin motor domain homology <MMO>

Query Match 95.3%; Score 5143; DB 2; Length 1028;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 989; Conservative 23; Mismatches 16; Indels 0; Gaps 0;

OY 17 MESALTRARVDGVODVLEENFTSEAFIENLRPRRENLYTYIGPLVSVNPRDLOI 76
I:|||||

```

```

DB 1 MSALTRARVDGVODVLEENFTSEAFIENLRPRRENLYTYIGPLVSVNPRDLOI 60
OY 77 YSRQHMERYRGVSFEYVPPHFLFAVDVYRALRTERDQAVMISSESGAGTEAKRLQ 136
DB 61 YSRQHMERYRGVSFEYVPPHFLFAVDVYRALRTERDQAVMISSESGAGTEAKRLQ 120
OY 137 FYAECPAPRGGAVDRLQSNPVEAFGNKTLRNDSSRFQKYMVDQDFKAPVGG 196
DB 121 FYAECPAPRGGAVDRLQSNPVEAFGNKTLRNDSSRFQKYMVDQDFKAPVGG 180
OY 197 HILSYLLEKSRVYHONHGERNFHVYOLLEGEDEETLRRLGLERNPOSTLYLVKQCAKY 256
DB 181 HILSYLLEKSRVYHONHGERNFHVYOLLEGEDEETLRRLGLERNPOSTLYLVKQCAKY 240
OY 257 SSINDSKDMKWKALSVDTFTEDEVEDLSTVASYLHGNHFADEDSNAQVTTENOL 316
DB 241 SSINDSKDMKWKALSVDTFTEDEVEDLSTVASYLHGNHFADEDSNAQVTTENOL 300
OY 317 KYTLRLGVEGTLRREALTHRKIIAKGEELLSPNLQAAVARDALAKAVSRTEFTWVR 376
DB 301 KYTLRLGVEGTLRREALTHRKIIAKGEELLSPNLQAAVARDALAKAVSRTEFTWVR 360
OY 377 KINRSLASKDAESPWSRSTVGLDLYGYEYVQHNSTFOFCINTYCNKLOOLEFTLTK 436
DB 361 KINRSLASKDAESPWSRSTVGLDLYGYEYVQHNSTFOFCINTYCNKLOOLEFTLTK 420
OY 437 SEOEYEAEGIAEPVOYFNNKIIICDLVEEKFKGIIISILDEECLRGEATDLTFLEKLED 496
DB 421 SEOEYEAEGIAEPVOYFNNKIIICDLVEEKFKGIIISILDEECLRGEATDLTFLEKLED 480
OY 497 TVKPHPHFLTHKLADOKTRKSLDRGEFRLHYAGEVTVYATGFLDKNNDLFFNLKETMC 556
DB 481 TVKPHPHFLTHKLADOKTRKSLDRGEFRLHYAGEVTVYATGFLDKNNDLFFNLKETMC 540
OY 557 SSNMPMAOCPSKSELSDDKRPETVATQFKMSLLQVLEILRSKPEATYIRCIKPNDAKOPG 616
DB 541 SSNMPMAOCPSKSELSDDKRPETVATQFKMSLLQVLEILRSKPEATYIRCIKPNDAKOPG 600
OY 617 RFDEVLIRHQQVLYGLMENLVRBRAGFAYRRKYEAFLQRYKSLCPEWPMAGRPDQGYA 676
DB 601 RFDEVLIRHQQVLYGLMENLVRBRAGFAYRRKYEAFLQRYKSLCPEWPMAGRPDQGYA 660
OY 677 VLVHRLGKPEEYKMGRTKIFIRPKTLFATEDSELEVRROSLATKIOAANGCFHMRQKFL 736
DB 661 VLVHRLGKPEEYKMGRTKIFIRPKTLFATEDSELEVRROSLATKIOAANGCFHMRQKFL 720
OY 737 RVKRSALCISQMMWGTIGRRKAARKWAQAOTIRRLINGFLRHSRCPENAFELDHVRAS 796
DB 721 RVKRSALCISQMMWGTIGRRKAARKWAQAOTIRRLINGFLRHSRCPENAFELDHVRAS 780
OY 797 FLNLRLRPLPNNVLDTSWPTPPALREASELLRELCKMNMWYKCRSISPEMKOOLQOKA 856
DB 781 FLNLRLRPLPNNVLDTSWPTPPALREASELLRELCKMNMWYKCRSISPEMKOOLQOKA 840
OY 857 VASELFFKCKNDYPOSVPRFLFSTRLGTEESIPRYLOSLGSEPIQYAVPVYKDYRKGYK 916
DB 841 VASELFFKCKNDYPOSVPRFLFSTRLGTEESIPRYLOSLGSEPIQYAVPVYKDYRKGYK 900
OY 917 RPRQLLTPSAVVIYEDAKVKQRIDYANLGTISVSSLSDSLFLVHVOQEDNKKQGDVYLO 976
DB 901 RSRQLLTPSAVVIYEDAKVKQRIDYANLGTISVSSLSDSLFLVHVOQEDNKKQGDVYLO 960
OY 977 SDHVETLTKTALSADRYNNININGSTTFAGGPGRGDIIDFTSGSELLITKAKNGHLAV 1036
DB 961 SDHVETLTKTALSADRYNNININGSTTFAGGPGRGDIIDFTSGSELLITKAKNGHLAV 1020
OY 1037 VAPRLNSR 1044
DB 1021 VAPRLNSR 1028

RESULT 3
S41749

```



```

17 MESALTRADRVGVODVLENTFTSEAFIENLRREPREMLITYIGPVLSVNPYRDLQI 76
1 MESALTRADRVGVODVLENTFTSEAFIENLRREPREMLITYIGPVLSVNPYRDLQI 60
77 YSRQHMERYKGVGFYEPVPHLEAVADTVYALRTERDQAVMISGSGAGKTEATKRLQ 136
61 YSRQHMERYKGVGFYEPVPHLEAVADTVYALRTERDQAVMISGSGAGKTEATKRLQ 120
137 FYAETCPAPRGAGVADRLQSNPVLFAFGNAKTLRNDSSRFKTYMDVQDFKGAPOVG 196
121 YVATCPGVSDQVEVYKRLQSNPVLFAFGNAKTLRNDSSRFKTYMDVQDFKGAPOVG 180
197 HILSYLLEKSRVYVHONGERNHFVYOLLEGGSEETLRIGLERNPQSTLYLKGGCAVY 256
181 HILSYLLEKSRVYVHONGERNHFVYOLLEGGSEETLRIGLERNPQSTLYLKGGCAVY 240
257 SSINDKSDMKVMKRALSVIDPTEDEVEDLSTIASVYLHGNHFADEDSNAQVTEENOL 316
241 SSINDKSDMKVMKRALSVIDPTEDEVEDLSTIASVYLHGNHFADEDSNAQVTEENOL 300
317 KYLRLLGVEGTTLRREALTRKRIIAKGEELISPLNEQAAYARDALAKAVYSRTFTWLYR 376
301 KYLRLLGVEGTTLRREALTRKRIIAKGEELISPLNEQAAYARDALAKAVYSRTFTWLYR 360
377 KINSLSKDAESPSMSTVYGLDLYGFEVPOHNSFEOPCINCYNEKTLQOLFIELTLK 436
361 KINSLSKDAESPSMSTVYGLDLYGFEVPOHNSFEOPCINCYNEKTLQOLFIELTLK 420
437 SEOEYEAEGIAMPVOYFNNKIICDLVEEKFGIISILDECLRPGEATDLFLEKLED 496
421 SEOEYEAEGIAMPVOYFNNKIICDLVEEKFGIISILDECLRPGEATDLFLEKLED 480
497 TVKPHPHLTKLADQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFLRLKETMC 556
481 TVKPHPHLTKLADQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFLRLKETMC 540
557 SSMPNIAOCDKSELSDKRRPETVATQFKMSLQLOVELILRSKPEAYIRICRPDAKOPG 616
541 DSGNFIACOFNRSLTDKRRPEATAQFKMSLQLOVELILRSKPEAYIRICRPDAKOPG 600
617 RFEVLIHQVYKYLGLMENLTVRRAGFAVRRKYEAFLOKYSKSLCPEPMWAGRPDQVA 676
601 RFEVLIHQVYKYLGLMENLTVRRAGFAVRRKYEAFLOKYSKSLCPEPMWAGRPDQVA 660
677 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGPFWROKFL 736
661 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGPFWROKFL 720
737 RVKRSALCIQSWMGRTGLRRKAARAKAAQITRLIRGFLIRHSRCPENAFELDHVYAS 796
721 HHKHSAAVEIQSWMGRTGLRRKAARAKAAQITRLIRGFLIRHSRCPENAFELDHVYAS 780
797 FLLNTRBQPRNVLDTSMPTRPPALREASELLEKCMNMWYKYSRSTISPEWKOLOQA 856
781 FLLNTRBQPRNVLDTSMPTRPPALREASELLEKCMNMWYKYSRSTISPEWKOLOQA 840
857 VASEIFKGGKKNYPOSVPRFLINTRLGNDEINTKILQOLESQTLTYAVPVVYKDKKYP 916
841 VASEIFKGGKKNYPOSVPRFLINTRLGNDEINTKILQOLESQTLTYAVPVVYKDKKYP 900
917 RPPOLLITPSAVVYVEDAKVORIDYANTLGISVSLSPSLFVTLVVOEDNKOQDVALY 976
901 RPPOLLITPSAVVYVEDAKVORIDYANTLGISVSLSPSLFVTLVVOEDNKOQDVALY 960
977 SDHVITLTKTALSDRVNININOSITFAGGPRDGIIDFTSGSELLITKAKNGHLAV 1036
961 SDHVITLTKTALSDRVNININOSITFAGGPRDGIIDFTSGSELLITKAKNGHLAV 1020
1037 VAPRLNSR 1044
1021 VAPRLNSR 1028

```

```

RESULT 5
H75634
myosin-1c - mouse (fragment)
N:Alternate names: myosin-I beta
C:Species: Mus musculus (house mouse)
C:date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 08-Sep-2000
C:Accession: H75634
R:crossref: F.; Amraoui, A.E.; Blanchard, S.; Lenoir, M.; Ripoll, C.; Vago, P.; Hamel,
Genomics 40, 332-341, 1997
A:title: Cloning of the genes encoding two murine and human cochlear unconventional t
A:Reference number: A59253; M01D:97237053; PMID:9119401
A:Accession: H75634
A:Status: preliminary; not compared with conceptual translation
A:molecule type: mRNA
A:Residues: 1-807 <CRO>
A:Cross-references: GB:X99638; NID:G1924960; PTDN:CA67956.1; PTD:G1924961
A:Experimental source: strain BALB/c; tissue type cochlea; dev stage adult
C:Genetics:
A:Gene: MGI:Myo1c
A:Cross-references: MGI:106612
A:Map position: 11:44.1
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
F;14-683/Domain: myosin motor domain homology <MMO>

```

```

Query Match 77.2% Score 4167; DB 2; Length 807;
Best local similarity 99.1%; Pred. No. 6, 3e-259; Indels 0; Gaps 0;
Matches 800; Conservative 3; Mismatches 4;
17 MESALTRADRVGVODVLENTFTSEAFIENLRREPREMLITYIGPVLSVNPYRDLQI 76
1 MESALTRADRVGVODVLENTFTSEAFIENLRREPREMLITYIGPVLSVNPYRDLQI 60
77 YSRQHMERYKGVGFYEPVPHLEAVADTVYALRTERDQAVMISGSGAGKTEATKRLQ 136
61 YSRQHMERYKGVGFYEPVPHLEAVADTVYALRTERDQAVMISGSGAGKTEATKRLQ 120
137 FYAETCPAPRGAGVADRLQSNPVLFAFGNAKTLRNDSSRFKTYMDVQDFKGAPOVG 196
121 FYAETCPAPRGAGVADRLQSNPVLFAFGNAKTLRNDSSRFKTYMDVQDFKGAPOVG 180
197 HILSYLLEKSRVYVHONGERNHFVYOLLEGGSEETLRIGLERNPQSTLYLKGGCAVY 256
181 HILSYLLEKSRVYVHONGERNHFVYOLLEGGSEETLRIGLERNPQSTLYLKGGCAVY 240
257 SSINDKSDMKVMKRALSVIDPTEDEVEDLSTIASVYLHGNHFADEDSNAQVTEENOL 316
241 SSINDKSDMKVMKRALSVIDPTEDEVEDLSTIASVYLHGNHFADEDSNAQVTEENOL 300
317 KYLRLLGVEGTTLRREALTRKRIIAKGEELISPLNEQAAYARDALAKAVYSRTFTWLYR 376
301 KYLRLLGVEGTTLRREALTRKRIIAKGEELISPLNEQAAYARDALAKAVYSRTFTWLYR 360
377 KINSLSKDAESPSMSTVYGLDLYGFEVPOHNSFEOPCINCYNEKTLQOLFIELTLK 436
361 KINSLSKDAESPSMSTVYGLDLYGFEVPOHNSFEOPCINCYNEKTLQOLFIELTLK 420
437 SEOEYEAEGIAMPVOYFNNKIICDLVEEKFGIISILDECLRPGEATDLFLEKLED 496
421 SEOEYEAEGIAMPVOYFNNKIICDLVEEKFGIISILDECLRPGEATDLFLEKLED 480
497 TVKPHPHLTKLADQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFLRLKETMC 556
481 TVKPHPHLTKLADQKTRKSLDGEFRLHYAGEVTVSYTGFLDKNNDLFLRLKETMC 540
557 SSMPNIAOCDKSELSDKRRPETVATQFKMSLQLOVELILRSKPEAYIRICRPDAKOPG 616
541 SSMPNIAOCDKSELSDKRRPETVATQFKMSLQLOVELILRSKPEAYIRICRPDAKOPG 600
617 RFEVLIHQVYKYLGLMENLTVRRAGFAVRRKYEAFLOKYSKSLCPEPMWAGRPDQVA 676
601 RFEVLIHQVYKYLGLMENLTVRRAGFAVRRKYEAFLOKYSKSLCPEPMWAGRPDQVA 660
677 VLVRLHGYKPEEYKMGRTKIFIRPPKTLFATEDSLEVRQSLATKIQANRGPFWROKFL 736

```



```

Db      661  VLVRLHLYGKPEEYKMGRTKIFRPPKTLFATEDSELEVRRSSTAKTIOAARGHGMQKEL 720
Oy      737  RKRSATCISQMWKSTLGRKRAAKRAAQAOTIRRLINGFLRHSRCPENAFLLDHYRAS 796
Db      721  RKRSATCISQMWKSTLGRKRAAKRAAQAOTIRRLINGFLRHSRCPENAFLLDHYRAS 780
Oy      797  FLNLRLRQLPRNVLDTSMPTRPALRE 823
Db      781  FLNLGRQLPRNVLDTSMPTRPALRD 807

```

RESULT 6

```

myosin IB - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Feb-2001
C:Accession: S45574
R:Strom-Morgan, N.; Skovronsky, D.M.; Artavanis-Tsakonas, S.; Mooseker, M.S.
J. Mol. Biol. 239, 347-356, 1994
A:Title: The molecular cloning and characterization of Drosophila melanogaster myosin-1A
A:Reference number: S45573; MUID:94260541; PMID:8201616
A:Accession: S45574
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1026 <STR>
A:Cross-references: GB:007596; NID:9466257; PIDN:AA19591.1; PID:9466258
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F:14-681/Domain: myosin motor domain homology <MOT>
F:105-112/Region: nucleotide-binding motif A (P-loop)

```

```

Query Match      48.8%; Score 2632; DB 2; Length 1026;
Best Local Similarity 51.5%; Pred. No. 2, 1e-160;
Matches 533; Conservative 177; Mismatches 300; Indels 24; Gaps 8;

```

```

Oy      17  MESALTPARVGVQDVLLLENFTSEAAFTENRRRRRENLITTYGGPVLVSNVPRDLOI 76
Db      1  METGLHEDRAGVQDVELLENYQSEAEATIGNLKRFOEDLITTYGVLISVNPYKOLPI 60
Oy      77  YSRQHMERYRGVSFEVPPHLPVADTVYRALTRRROAVVISESGAGTTEATKRLQ 136
Db      61  YTDNHKAVKRNHFEMPHIFAVVDNAFRSLIEENKQCVLISESSSGKTEASKVQLQ 120
Oy      137  FYAETCPARERGAARDLLQSNPVLEAFGNAKTLIRNDSSRFKGYMDQDFGAPYVG 196
Db      121  FIAACGNOTVEGVKDKLNSPVLEAFGNAKTLIRNDSSRFKGYMDQDFGAPYVG 180
Oy      197  HILSYLLKSRVYVHONHGERNFHVYOLLEGGEETLRLGLERNPOSYLLVYGQCAKY 256
Db      181  NILNTLLKSRVYVHONHGERNFHVYOLLEGGEETLRLGLERNPOSYLLVYGQCAKY 240
Oy      257  SSINDKSDMKVKKALSVYIDETEDVLDLSTVSLHGINHFADEDSNAQVTTENOL 316
Db      241  TSINDADSEFKVQOALTYIDEGKEQREIFGIVASLHLGVNGV-TEVEGNAKANSDLY 299
Oy      317  KVLTRLLRGEGTTLREALTHKRIIAKGELLSPNLLEDAVAAARALAKAVYSRTTWLVR 376
Db      300  VTAAILLVNASELALTHKRIIDARGVVTSPLNOELAIYARALAKAVYDRLFLSWLQ 359
Oy      377  KINRSIASDAESPMSRSTVGLDIDYGFVFOHNSPEOCINCYCEKILQOLFETLTK 436
Db      360  RLNISLQAKETRA---SRNNVMGILIDYGFIFOKNSPEOCINFCNEKILQOLFETLTK 416
Oy      437  SPOEYEAAGTAMEVOYFNKKITCDLVEEKFGIISILDECLRPGEATDLTFLEKLD 496
Db      417  SPQDYRRREGLEWLEVEYFNKVICNLLEEKHKGIIISILDECLRPGEATDLTFLEKLD 476
Oy      497  TVKPRHFLTKLADQKTRKSLDGEFFLLHAGEVYVSYTGFLDKNDLFLRKLKTM 556
Db      477  KLAQHHYVYCHEKAPAHIKKIMLDERFLVHAYAGEVYVSYNGFLDKNDLFLRDLKETS 536
Oy      557  SSMPNIMAOCEKSELSDKKRPETVANOQFKMSLLQVLEILRSKEPAVIRICIKPNDAKOPG 616

```

```

Db      537  KAGNGIVASCPEPEKELSLKRPETAITQFRASLNNLMIDLDEPSTIRIKRNDLQITAN 596
Oy      617  REDEVLLRHQVKTGYLGMENLRVRAGATYRRKKEAFLOKRSKLCPETPWM---AGRPDQ 674
Db      597  VFNDLVHQQVKTGYLGMENLRVRAGATYRRKKEAFLOKRSKLCPETPWM---AGRPDQ 656
Oy      675  VAVLVRHLGYKPEEYKMGRTKIFRPPKTLFATEDSELEVRRSSTAKTIOAARGHGMQKEL 734
Db      657  VQVLKDLGMDPEEKYRVGTEKFLRMPRTLFOTEDYAOEKKEHIAIIOHMGIMQKRR 716
Oy      735  FLVRKSAICISQMWKSTLGRKRAAKRAAQAOTIRRLINGFLRHSRCPENAFLLDHYR 794
Db      717  YLTKRAOYIIMQSCRRKLAQQAARAKREAAKIRAFIKGFTIRNDAPNGFDEETANX 776
Oy      795  ASFLNLRLRQLPRNVLDTSMPTRPALREASELLRCLKNMVKCRISPEKQOIQ 854
Db      777  RMWLLRLAKELPKVLDKSPHAPRGCEASGLHRLHLRLARITRLTLPOOKROFEL 836
Oy      855  KAVASELFGKKKNDYQSVPRLEISTRLGTEEI---SPVLOSLSGEPIQVAVPVYKYDR 911
Db      837  KVLAEKVFKKKNNYASSVSTWFQEDRIPREHIQVNDPVASFQSEDLKYSGFCTKDR 896
Oy      912  KGKPRPROLLRPSANVYVEDAKVQ-----RIDVANLGISVSSLSDSLFLVAVQR 964
Db      897  HGKSRDRFTLLSKNATYVLDGTYRQKHRLPLDKIDF-----FLTNHNDLAWIRIPL 950
Oy      965  EDNRKQDVLQSDHYETLTKTALSADRVNINI-NQSIYFAGGPGDIDTTSQSE 1023
Db      951  DLKKDGDLLIPLRIEFTSTYIDIVGASIVSYDRNSLHNHYVKGAGVIDIQTGA 1010
Oy      1024  LLITTKAKNGHLAV 1037
Db      1011  PGVVRDK-GHLVIT 1023

```

RESULT 7

```

B45438
myosin I beta, Mxi beta - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 31-Jan-2000
C:Accession: B45438
R:Sherr, E.H.; Joyce, M.P.; Greene, L.A.
J. Cell Biol. 120, 1405-1416, 1993
A:Title: Mammalian myosin I alpha, I beta, and I gamma: new widely expressed genes of
A:Reference number: A45436; MUID:93194946; PMID:8445986
A:Accession: B45438
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-448 <SHE>
A:Note: sequence extracted from NCBI backbone (NCBI:131911)
C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
F:1-448/Domain: myosin motor domain homology (fragment) <MOT>

```

```

Query Match      41.7%; Score 2252; DB 2; Length 448;
Best Local Similarity 97.1%; Pred. No. 1, 3e-136;
Matches 435; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy      162  LEAFGNATLRNDSSRFGKTYMDVDFKAPVGGHILSYLLKSRVYVHONHGERNFHV 221
Db      1  LEAFGNATLRNDSSRFGKTYMDVDFKAPVGGHILSYLLKSRVYVHONHGERNFHV 60
Oy      222  YOLLEGGEETLRLGLERNPOSYLLVYGQCAVSSINDKSDMKVKKALSVYIDETED 281
Db      61  YOLLEGGEETLRLGLERNPOSYLLVYGQCAVSSINDKSDMKVKKALSVYIDETED 120
Oy      282  VEDLLSYAVSLHGINHFADEDSNAQVTTENOLKYLTRLLGEGTTLREALTHKRIIA 341
Db      121  VEDLLSYAVSLHGINHFADEDSNAQVTTENOLKYLTRLLGEGTTLREALTHKRIIA 180
Oy      342  KGEELSPNLNLEDAVAAARALAKAVYSRTTWLVRKINRSIASDAESPMSRSTVGL 401
Db      181  KGEELSPNLNLEDAVAAARALAKAVYSRTTWLVRKINRSIASDAESPMSRSTVGL 240

```

402 DYGEVEYOHNSFEOFCINCNKELQOLFEITLSEKSEOEYENAGIAMEPVQYFNKIK 461
 462 DYGEVEYOHNSFEOFCINCNKELQOLFEITLSEKSEOEYENAGIAMEPVQYFNKIK 300
 241 DYGEVEYOHNSFEOFCINCNKELQOLFEITLSEKSEOEYENAGIAMEPVQYFNKIK 300
 462 DYGEVEYOHNSFEOFCINCNKELQOLFEITLSEKSEOEYENAGIAMEPVQYFNKIK 521
 301 DYGEVEYOHNSFEOFCINCNKELQOLFEITLSEKSEOEYENAGIAMEPVQYFNKIK 360
 522 EFRLLHYAGEVTSVTFGLDKNNDLFRNLKETSMSMNPIMAOCEFSKSELSDDKRPETV 581
 361 EFRLLHYAGEVTSVTFGLDKNNDLFRNLKETSMSMNPIMAOCEFSKSELSDDKRPETV 420
 582 ATOFMSLLOLVELIRSKPEAYTRICRP 609
 421 ATOFMSLLOLVELIRSKPEAYTRICRP 448
 RESULT 8
 B45439
 myosin-I, Myr 1c (alternatively spliced) - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
 C:Accession: B45439
 R:Ruppert, C.; Kroschewski, R.; Bahler, M.
 J. Cell Biol. 120, 1393-1403, 1993
 A:Title: Identification, characterization and cloning of myr 1, a mammalian myosin-I.
 A:Reference number: A45439; MUID:93194945; PMID:8449985
 A:Accession: B45439
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1078 <RUP>
 A:Note: Sequence extracted from NCBI backbone (NCBI:128956, NCBI:128986)
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: nucleotide binding; P-loop
 F:18-688/Domain: myosin motor domain homology <MMOT>
 F:108-115/Region: nucleotide-binding motif A (P-loop)
 Query Match 39.1%; Score 2109; DB 2; Length 1078;
 Best local similarity 42.8%; Pred. No. 7.5e-127;
 Matches 463; Conservative 183; Mismatches 346; Indels 90; Gaps 19;
 27 VGVDFVLLNENFSEAFENLRRFRNLIYIYIGVAVSNVPRLOIYRQHMERYR 86
 15 IGVGDTVLEPL-NEETFDLNKRRFNEIYIYIGVAVSNVPRSLPIYSPKVEDYR 73
 87 GVSFEVPPHLEAVADYVRAIRTERDOAVMISGSGAGTETATKRLIOFYAETCPAPE 146
 74 NRRFELSPHIFALSDENVRSLRDOCKDQCLITGESGAGTETASKLWSTVAVACGKA 133
 147 RGVAVDRLLQSNPVLEAFGNKATLRNNSRFGKMDVQDFGAPVGHILSYLLEKS 206
 134 EVNQVKEQLQSTPVLEAFGNKATLRNNSRFGKMDVQDFGAPVGHILSYLLEKS 193
 207 RVVHONHGERNFHVFYQLIEGSEETLRRLGLERNPQSYLYLVKQCAKVSINDKSMK 266
 194 RVVQCPGRGRNFHVFYQLIEGSEETLRRLGLERNPQSYLYLVKQCAKVSINDKSMK 252
 267 VMKALSVIDFTEDEEDLSTVAVLHGNHFAADSDN- - -AQTNNQQLYTLRL 322
 253 TVNNAMQIVGFSPESESVLEVAVALKGINTEFEPESRBMGLDESKIKDNEKEICEL 312
 323 LGVEGTTLEALTHRRKIKAQBELSPLNLEQAAVARDALAKAVYSRTFTWLVKRNLSL 382
 313 TSDQVVLERRASFRRIVEAKQKVSFTTLNVAQAYARDAKNTLSRFLSWLVNKNESI 372
 383 ASKDAESPSTRTTVGLDIYGEVFOHNSFEOFCINCNKELQOLFEITLSEKSEOEY 442
 373 KAQTKVKK- - -VGVLDIYGEVFOHNSFEOFCINCNKELQOLFEITLSEKSEOEY 427
 443 EAAGIAMEPVQYFNKIKDYGEVEYOHNSFEOFCINCNKELQOLFEITLSEKSEOEY 502
 428 IREDIETHYHIDYNNAILICDLIENNNIGLAMDCECLARPGVYVTEFLEKLNQVCATHO 487

503 HFLTH- -KLADQTRKSLDGEFRILHYAGEVTSVTFGLDKNNDLFRNLKETSMSMNP 560
 488 HFESRMSKSRPLNDTTLPSCHRIQHYAGVTSVTFGLDKNNDLFRNLKETSMSMNP 547
 561 PIMAOCEFSKSELS- -DKRPEVATQFMSLLOLVELIRSKPEAYTRICRPDAKOPGRF 618
 548 ALISLPEPENGPAKVNLRPTAGSQFASVATAMNLTQKPNYTRICRPDAKOPGRF 607
 619 DEVILIRQVYKLGIMENLRVRAGFAVRKRYEAFLORYSLDPEYPMAGRPQDGVAVL 678
 608 SESLVCHQIRYLGILENVRVRAGFAVRKRYEAFLORYSLDPEYPMAGRPQDGVAVL 667
 679 VRHGYPRPEYKMGRTKIFRFPKTLFATEDSLEVRQSLATKIQAMGFMHOKFELV 738
 668 FNELEIPEVEEYSPGRSKIFIRNPRTLEQLDELDKORLEDLATLQKTYGMRCKRHFLLM 727
 739 KRSALCISQSWRGTLGRKAKKAKKMAQOTIRLRGE- - -ILRHSRPEEN- - - 787
 728 KRSQVIAWYRRYRRAQQRKRYQKSSALVQSYIRGRKAKKILRELKHOKRCEATTTA 787
 788 - - - - -PELDHVRASFLLNLRPLR- - -NVLDTSMPTPPA 820
 788 AYHGTQVRYREYRKRFRANAGKTYETTLQRIYQKYLEKKNMPSLSPIDKNWPSRPLY 847
 821 LREASELRELCKRMWV- - -KYCRSISPEKQOOLQKAVASELFGKADNTPQSYRLE 877
 848 FLDSH- - -KELKRIFLHLMRCKKRYDQFTDQKLYEKELEASLEFDKALVPSVSGQFP 905
 878 ISTRGTSEIIP- - -RYLQSLGSEPIQVAVPVYKRYKRPRLQTLTPSAVIVE- 932
 906 OGAVL- - -EINKNPKYKALKAIEKIIIAEVVAKIRANAGKISRIFLLNNLLADQ 962
 933 -DAKYKORIDYANLIGVSSLSLSEVLAVQV-EDNKQGDVAVLQSDHYETLTK- - -T 987
 963 KSGQKSEVPLVDYVAVKMSQNDGFVAFHLEKESSEASGDFLFSDDLIEATVLYRI 1022
 988 ALSADRVN-NINI- - -NGSITTEGPGRGDITDFTSGSELLTKKNGHLAV 1036
 1023 TISQTKQKLNIEISDEFLVQRPQDVKVFIQGNQKNSVP- - - - -TCKKNNRLL 1074
 1037 VA 1038
 1075 VA 1076
 RESULT 9
 A29483
 myosin heavy chain I, brush border - bovine
 N/Alternate names: myosin I heavy chain-like protein, MIRC
 N/Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
 C:Accession: A29483; JX0208
 R:Hoshimaru, M.; Nakaniishi, S.
 J. Biol. Chem. 262, 14625-14632, 1987
 A:Title: Identification of a new type of mammalian myosin heavy chain by molecular cl
 A:Reference number: A29483; MUID:86033016; PMID:3667594
 A:Accession: A29483
 A:Molecule type: mRNA
 A:Residues: 1-1043 <HOS>
 A:Cross-references: GB:J02819; NID:9163405; PIDN:AAA30658.1; PID:9163406
 R:Kawakami, H.; Moriyoshi, K.; Utsun, T.; Nakaniishi, S.
 J. Biochem. 111, 302-309, 1992
 A:Title: Structural organization and expression of the gene for bovine myosin I heavy
 A:Reference number: JX0208; MUID:92268028; PMID:1587791
 A:Accession: JX0208
 A:Molecule type: DNA
 A:Residues: 1-789 <RAW>
 A:Note: nucleotide sequence is not complete
 C:Genetics: 38/3; 77/2; 109/1; 144/1; 159/3; 181/1; 214/1; 248/3; 298/1; 337/3; 366/3;
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology

C:Keywords: actin binding; hydrolase; intestine; nucleotide binding; p-loop
 F:11-681/Domain: myosin motor domain homology <MMO>
 F:101-108/Region: nucleotide-binding motif A (p-loop)
 F:572-593/Region: actin binding #status predicted
 F:723-1043/Domain: carboxyl-terminal <CTD>
 F:107/Binding site: ATP (Lys) #status predicted

Query Match 39.0%; Score 2106.5; DB 1; Length 1043;
 Best Local Similarity 44.0%; Pred. No. 1e-126;
 Matches 455; Conservative 185; Mismatches 330; Indels 65; Gaps 21;

```

27  VGVDDVLLNENFSEAAFIENLRRELRRLTYIGPVLYSVNPRDQIYSRQHMYR 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8  VGVDDVLLNLEPDE-SLRNQLQRYKKETIYIGVLYSVNPRDQIYIDLEFAKRR 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87  GVSFEVPPHLEAVADVRLATERDDAVMSGGSGAGKTEATRLQFAETCPAPE 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67  DYTTELKPHITVLANNAVQLDRDRDQILITGSSGAGKTEASLVSYAAVCGKGE 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147  RGAVARDRLLQSNPVLEAFGNATKLIRNDSSRGKMYDQFDPKAPVGGHLSYLEKS 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127  QVNSVKEQLQSNPVLEAFGNATKLIRNDSSRGKMYDQFDPKAPVGGHLSYLEKS 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207  RVVQNHGERNFHFVYQLLEGEEFLRLGLERNPQSYLYLVKGOCARVSSINDKSPDK 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187  RVVQNHGERNFHFVYQLLEGEEFLRLGLERNPQSYLYLVKGOCARVSSINDKSPDK 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267  VMKRALSVDFTEDEVEDLSTVAVLHGNHFADEDSN----AQVTENQLKYTRL 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
246  VLQSNAMTVIGFSDDELRYQVLEAVLVKLGWELNEFQANQVPAISGIDGVOEISEL 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323  LGEVGTTLREALTHRKIIAKGELLSPNLEQAAVARDALAKAVYSRTFTWLVRKINSL 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306  VGLNVELERLALCSRMETAKERYVTTLVNIOAQVARDALAKAVYSRTFTWLVRKINSL 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383  ASKDAESPWSRSTTVGLDIYGEVFOHNSFEOCINCKLQOULETLTKSEOEY 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
366  KVGTGE-----KKVWGVLDIYGEFLEEDNSEFOYINCNKLOQVETEMTLKEOEY 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443  EAEGIMEPVOYFNKKIICDLVEEKFGIISILDECLRPGEATD/LFLEKLEDPVKRHP 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421  KRGIQVWVYVEYDNGIICNLEHNGRILAMDECLRPGVSDSTFLAKINOLFSAHS 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503  HFLTHLADQKTRK---SLDRGEFLRLHAGVTVSYVGLDKNNDLFRMLKETGSSM 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481  HY-ESVYTONAQRODHSMGLSCFICHYAGKVTVNNSFLDKNNDLFRLSOAMMKAR 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
560  NPIMACDFKSE--LSDKRPETVATQFKMSLQVETLRSEKPAYICIRPNDAKQGR 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
540  HPLRLSLFPEDGPKQASLKRPTAGQKSSVTTLMKMLYSKNPVIICIRPNEHQGRH 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
618  FDEVLIRHVKYTLGLMENVLRVRAQFAYRRKYEAFLQRYKSLCEPTWPMAGRPDQVAV 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
600  FSEFLVSVQAYLGLLENVRVRAQFAYRRKYEAFLQRYKSLCEPTWPMAGRPDQVAV 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
678  LVRLHLYGKDEEYKMGRTKIFIRPKTLFATEDSLEVRQSALTIQOAMRFFHRRKQFLR 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
660  VIGELSMSEELAFKTKITFIRPKTLFATEDSLEVRQSALTIQOAMRFFHRRKQFLR 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738  VRSALCIQSNMGRGLGRRKAKRKAQAOTIRRLRGLRHSR-----CPENAF 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
720  MKRSOIVISWFRGNMCKHYRKMAKASALLIOAFRGWKARKNYRKFRSAGALLISNFI 779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
789  FLDAHVRASFLNLRQLPR-NVLDTSMPRP-----PALRASLDELCKKNNVYCR 842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
780  YKSMQV-KFLGLKNDLPSPSTLIDKMPKPSADYKFNANHLQRLFRQ-----MK-CK 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
843  ----SISPMKQOLQOKAVASEIFKGGKDNYPQSVRLSTIRGTEISPRVLYSGSE 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
831  KRPDLSPKROYEVLREKLCASELFGKKASYPQSVRLSTIRGTEISPRVLYSGSE 889
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
899  --PIQYAVPVVYKDKKGRPRRDLTPSAVVI--VEDAKVQRIDYANITGISVSLIS 954
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

DB 890  EGPILMAETVYKVRGNKAKTSSRILLTKGHVITTDKRNPKQAKVPIPLNSIAGVTSRK 949
OY 955  DSLFVLAHQREDN-KQKGDVYDSDHYIELTK---TALSADVN-----NINIQ 1001
DB 950  DGLSHLSSEISSGSGGEELVSEHVELTTLICRATLDATQMLPVYTEEFSEVFKR 1009
OY 1002  GSTI-PAGEGRDG 1014
DB 1010  GSLTVKVIQGPGGG 1024
  
```

RESUT 10

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
 C:Accession: C45439
 J:Ruppert, C.; Kroschewski, R.; Bahler, M.
 J:Cell Biol. 120, 1393-1403, 1993
 A:Title: Identification, characterization and cloning of myr 1, a mammalian myosin-I.
 A:Reference number: A45439; MUID:93194945; PMID:8449985
 A:Accession: C45439
 A:Status: preliminary
 A:Molecule type: mRNA; Protein
 A:Residues: 1-1107 <RUP>
 A:Note: sequence extracted from NCBI backbone (NCBI:128956, NCBI:128984)
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: nucleotide binding; p-loop
 F:108-115/Region: nucleotide-binding motif A (p-loop)

Query Match 38.8%; Score 2094.5; DB 2; Length 1107;
 Best Local Similarity 41.7%; Pred. No. 6.6e-126;
 Matches 463; Conservative 163; Mismatches 346; Indels 119; Gaps 19;

```

27  VGVDDVLLNENFSEAAFIENLRRELRRLTYIGPVLYSVNPRDQIYSRQHMYR 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
15  IGVGDVLLLEPL-NETFLDNKKRKHEDHETIYIGVLYSVNPRDQIYSRQHMYR 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87  GVSFEVPPHLEAVADVRLATERDDAVMSGGSGAGKTEATRLQFAETCPAPE 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
74  NRVFELSPHIALSDQEARSLRDQKDCILITGSSGAGKTEASLVSYAAVCGKGA 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147  RGAVARDRLLQSNPVLEAFGNATKLIRNDSSRGKMYDQFDPKAPVGGHLSYLEKS 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134  EYVQVKEQLQSTPVLEAFGNATKLIRNDSSRGKMYDQFDPKAPVGGHLSYLEKS 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207  RVVQNHGERNFHFVYQLLEGEEFLRLGLERNPQSYLYLVKGOCARVSSINDKSPDK 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
194  RVVQNHGERNFHFVYQLLEGEEFLRLGLERNPQSYLYLVKGOCARVSSINDKSPDK 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267  VMKRALSVDFTEDEVEDLSTVAVLHGNHFADEDSN----AQVTENQLKYTRL 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253  TVRNAQIVGFSDDPAESVLEVAVALKLGNEEFPEPSRMNGLDSEKIKDNEKEICEL 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
323  LGEVGTTLREALTHRKIIAKGELLSPNLEQAAVARDALAKAVYSRTFTWLVRKINSL 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
313  TSIDQVLERARSFPTVEAKQKYSTIYLANQATYARDALKKMLYSRFSWLVNRINESI 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
383  ASKDAESPWSRSTTVGLDIYGEVFOHNSFEOCINCKLQOULETLTKSEOEY 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
373  KQOTVVRK-----VWGVLDIYGEFLEEDNSEFOYINCNKLOQVETEMTLKEOEY 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443  EAEGIMEPVOYFNKKIICDLVEEKFGIISILDECLRPGEATD/LFLEKLEDPVKRHP 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428  IREDLEWHIDYFNNAIICDLLENNTGICLAMDDECLRPQVTVDEFELKLNQVCAHQ 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503  HFLTHLADQKTRKSLDRGEFLRLHAGVTVSYVGLDKNNDLFRMLKETGSSM 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
488  HESRMSKCSFRLNDTTLPHSCFRIQYAGKVLYQVGGVFDKNNDLFRLSOAMMKAR 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
561  PIMACDFKSELS--DKRPETVATQFKMSLQVETLRSEKPAYICIRPNDAKQGR 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

Db 548 ALIKSLPEGNPAKVNLRKPPYAGSQCOKASVAILMKVLQTKNPNYINCIRPNCKKAHIF 607
 Qy 619 DEVLIRHQVKYLGIMENLRVRAGFAVRKRYEAFLOQKSLCPEWPMAGRPDQVAVL 678
 Db 608 SESLVCHQIRVLGLLENVRBAGYAFROAEPCELEKMKLCQTMWPMKGPASGVEVL 667
 Qy 679 VRLHGYRPEEKMKRTKIFIRPFTLPADESLVRQSLATKIQAMRPFHMQRKFLRV 738
 Db 668 FNELEIPVEEYSFGRSKIFIRNPRTLFOLEDRLQRLEDLTLQKTLRGKCKCTHFLM 727
 Qy 739 KRSALCISQWVRGTLGRKRAKRAKMAOTIRRLIRGF-----ILKHSPPCPENA---- 787
 Db 728 KRSQVYLAAMVRRRAQOKRYOIKSSALVIOSTYIRGMKARKILBELKHOKRCKEAAVTIA 787
 Qy 788 -----FPLDH 792
 Db 788 AYMHGTQARRELKRLKEBARRRHAAVAVIWAAYMLGLKYRREYRKFEPANAGKIYEFTLQR 847
 Qy 793 VRASFLMLRROLPR-NVLDTSWPTPPALREASELRELCKMKNVW---KYCRSISPEW 848
 Db 848 IVQKYLEMKMKMPSLSPIDKMPSRPLFLDSTH--KELKRIFHLMCKKRYRQGFDDQ 905
 Qy 849 KOOLQOKAVASEIFKGGKNDYPOSVPLFLISTRLGTEBISP---RVLOSIGSEPIQYAV 904
 Db 906 KLIVEKLEASLFEKDKALYPSVQCPQAYL---EINKNPYKKIKLDAIEEKIIIAE 962
 Qy 905 PYYKIDRGKYPRLQLLTSPSAVIVF--DAKYQORIDYANLIGISVSLSDSLFVHV 962
 Db 963 VYNNKINRANGSTSTIFLITNNLLLAQOKSGQIKSEVPLVDYTKVSSQNDGFFAVHL 1022
 Qy 963 QR-EDNKOQGDVVLQSDVHVIETLTK--TALSADRVN-NINI-----NCGSITFA 1007
 Db 1023 KGSSEAKSGDGLFSDHLEMAKTLKYLRTLSQTKQKINIEISDEFLVOFQDKVCYKFI 1082
 Qy 1008 GSGPGDGIIDTSSGSELLITKAKNGHLAVVA 1038
 Db 1083 QGNOKNGSVP-----TCRKNRNLLEVA 1105

RESULT 11

A45438
 myosin I alpha chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
 C:Accession: A45438
 J. Cell Biol. 120, 1405-1416, 1993
 A:Title: Mammalian myosin I alpha, I beta, and I gamma: new widely expressed genes of the
 A:Reference number: A45438; M0ID:93194946; PMID:8449986
 A:Accession: A45438
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-1079 <SHE>
 A:Cross-references: GB:100923; NID:9200008; PIDN:AAA39800.1; PID:9200009
 A:Experimental source: brain
 A:Note: sequence extracted from NCBI backbone (NCBI:129047, NCBI:129048)
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: nucleotide binding; P-loop
 F:18-689/Domain: myosin motor domain homology <MMOT>
 F:108-115/Region: nucleotide-binding motif A (P-loop)
 Query Match 38.6%; Score 2081.5; DB 2; Length 1079;
 Best Local Similarity 42.7%; Pred. No. 4.3e-125;
 Matches 463; Conservative 181; Mismatches 347; Indels 93; Gaps 20;

Qy 27 VGVODFVLENFSEAFLENLRRENLIYIYIGPVSVNYPRLQIYRQHMERYR 86
 Db 15 IGVGDVNLLEPL-NEETFLDNLKKRFDHNEIYIYISVNPYKSLTIYSEKEDYR 73
 Qy 87 GVSFEYVPHLFAVDYVRAALTEERRDQAVMISGSGAGTEATRLQIYAEVTCPAPE 146
 Db 74 NRPYELSPHIFALSEAVRSLRDQKDCILITGSGAGKTEASKLMSVYAAVCGKA 133

Qy 147 RGVAVDRLLQSNPVLEAFGNKATLINDNSSRFGKMYDVOFEGKAPVGGHILSYLEKS 206
 Db 134 EVNQVKEQLQSNPVLEAFGNKATYVNDNSSRFGKMYDVOFEGKAPVGGHILSYLEKS 193
 Qy 207 RVVHONHGERBNHVFQLLGEGEELRLGLGERNPQSLYLKVOQCAKVSINDKSPK 266
 Db 194 RVKQCPGRGERBNHVFQLLGEGEELRLGLGERNPQSLYLKVOQCAKVSINDKSPK 252
 Qy 267 VMKRLSVIDEFDEVEDLLSIYASVILGNIFHAD-----EDSNAQVTTENQKLYL 320
 Db 253 TVRNAMQYGFEDHEAEVLEVAVALKGNIEFKRESVNGIDESKIKDKIELNEKFAV 312
 Qy 321 RLIVGEGTLLREALTRKTIANGELISPLNEQAAADALAKAVYSTFTWLYRKINR 380
 Db 313 RPASVK-VVLERAFSEFVEAKREKYSTLNVAAQAYADALAKINISFLMVLNRIWE 371
 Qy 381 SLASKDAESPSMSTVIGLIDYGEVPHONSFEQFCINCEKLOQFIELTLKSEOE 440
 Db 372 SIAQKIVKRRK-----VMGVLDYGFIEIEDNSFEPFIINCEKLOQFIELTLKSEOE 426
 Qy 441 EYEAEGIAWEPVOYFNKKIICDLVEKFKGIIISIDECELRGEATDLFLEKLEDYVP 500
 Db 427 EYIREDIEWHIDYFNNAIICDLINNTNGIILAMDECELRGEATDLFLEKLEDYVP 486
 Qy 501 HPHFLTH--KLADQKTKSLDRGEFRLHAYAGEVTVSYTGDKNNDLFRLNKEFMCS 558
 Db 487 HOHEESMKSCKSRLNDTLPHSCFRQIHTAGVVLQVBSGVKNDLILYRDSQAMKA 546
 Qy 559 NNPIMACCFKSELS--DKRREPVAQFMKSLLOVLEILRSKPEAYIRICIRPNDAKOPG 616
 Db 547 DHSILKSLFPEGNAKAKNLRPPTAGSQFASVATLBRNLOTKNPNYIRICIRPNDAK 606
 Qy 617 RFDEVLIRHQVKYLGIMENLRVRAGFAVRKRYEAFLOQKSLCPEWPMAGRPDQVAVL 676
 Db 607 IFNESLVCHQIRVLGLLENVRBAGYAFROAEPCELEKMKLCQTMWPMKGPASGVEVL 666
 Qy 677 VVRLHGYRPEEKMKRTKIFIRPFTLPADESLVRQSLATKIQAMRPFHMQRKFLRV 736
 Db 667 FNELEIPVEEYSFGRSKIFIRNPRTLFOLEDRLQRLEDLTLQKTLRGKCKCTHFLM 726
 Qy 737 KRSALCISQWVRGTLGRKRAKRAKMAOTIRRLIRGF-----ILKHSPPCPENA-- 787
 Db 727 KRSQVYLAAMVRRRAQOKRYOIKSSALVIOSTYIRGMKARKILBELKHOKRCKEAAVTIA 786
 Qy 788 -----FPLDH 792
 Db 788 AYMHGTQARRELKRLKEBARRRHAAVAVIWAAYMLGLKYRREYRKFEPANAGKIYEFTLQR 847
 Qy 793 VRASFLMLRROLPR-NVLDTSWPTPPALREASELRELCKMKNVW---KYCRSISPEW 848
 Db 848 IVQKYLEMKMKMPSLSPIDKMPSRPLFLDSTH--KELKRIFHLMCKKRYRQGFDDQ 905
 Qy 849 KOOLQOKAVASEIFKGGKNDYPOSVPLFLISTRLGTEBISP---RVLOSIGSEPIQYAV 904
 Db 906 KLIVEKLEASLFEKDKALYPSVQCPQAYL---EINKNPYKKIKLDAIEEKIIIAE 962
 Qy 905 PYYKIDRGKYPRLQLLTSPSAVIVF--DAKYQORIDYANLIGISVSLSDSLFVHV 962
 Db 963 VYNNKINRANGSTSTIFLITNNLLLAQOKSGQIKSEVPLVDYTKVSSQNDGFFAVHL 1022
 Qy 963 QR-EDNKOQGDVVLQSDVHVIETLTK--TALSADRVN-NINI-----NCGSITFA 1007
 Db 1023 KGSSEAKSGDGLFSDHLEMAKTLKYLRTLSQTKQKINIEISDEFLVOFQDKVCYKFI 1082
 Qy 1008 GSGPGDGIIDTSSGSELLITKAKNGHLAVVA 1038
 Db 1083 QGNOKNGSVP-----TCRKNRNLLEVA 1105

RESULT 12

A45439
 myosin I heavy chain - rat
 C:Species: Rattus norvegicus (Norway rat)

Qy 1022 RTLSQTKQKINIEISDEFLVOFQDKVCYKFIQGNOKNGSVP-----TCRKNRNL 1073
 Db 1035 AVVA 1038
 Db 1074 LEVA 1077

C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
 C:Accession: A45439; S29984
 R:Ruppert, C.; Kroschewski, R.; Bahler, M.
 J. Cell Biol. 120, 1393-1403, 1993
 A:Title: Identification, characterization and cloning of myr 1, a mammalian myosin-I.
 A:Reference number: A45439; MUID:93194945; PMID:8449985
 A:Accession: A45439
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1136 <RUP>
 A:Cross-References: GB:X68199; NID:956732; PIDN:CAA48287.1; PID:956733
 A:Note: Sequence extracted from NCBI backbone (NCBIN:128956, NCBI:128961)
 R:Ruppert, C.
 submitted to the EMBL Data Library, September 1992
 A:Reference number: S29984
 A:Accession: S29984
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1136 <R02>
 A:Cross-References: EMBL:X68199; NID:956732; PIDN:CAA48287.1; PID:956733
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: nucleotide binding; P-loop
 F:18-688/Domain: myosin motor domain homology <MOT>
 F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 38.5%; Score 2080; DB 2; Length 1136;
 Best Local Similarity 40.6%; Pred. No. 5.9e-125;
 Matches 463; Conservative 183; Mismatches 346; Indels 148; Gaps 19;

27 VGVQDPVLENTSEAFIENLRRENLYITGVPVSVNPDQIYSRQHMERYR 86
 15 IGVGTVLEPL-NEETFDNLKRRFDHNEYITGYSVIVSNPYSRSLPIYSPEKVEDYR 73
 87 GVSFEVPEPLFAVDVTVRALTERRDOAVMISEGAGTEAKRRLQFAETCPAPE 146
 74 NNFTELSPHITALDEATRSRLDQDQOCILITGESAGKTEKSLVMSVAAYCGKA 133
 147 RGVAVRDLRLSNPVLFAFNKATLRNDSSRFKYMVDQDFKAPVCGHILSYLTKS 206
 134 EVNQVKEQLQSTPLFAFNKATLRNDSSRFKYMVDQDFKAPVCGHILSYLTKS 193
 207 RYVHNGEERNHYVYQLLEGGEETLRKLERPOSTLYLVKQCAKVSINDKDMK 266
 134 RYVHNGEERNHYVYQLLEGGEETLRKLERPOSTLYLVKQCAKVSINDKDMK 266
 134 RYVHNGEERNHYVYQLLEGGEETLRKLERPOSTLYLVKQCAKVSINDKDMK 266
 267 VMKALSYIDFTEDEVEDLSTIVASVLIHGNTHFAADSDS---AQTENOLKYLRL 322
 253 TVRNAMQIVGSDPAESVLEVAVALKGNIEKRPESRMAGLDESKIKDKNELKEICEL 312
 323 LGVEETLRREALTHRKITAKEEELLPLNLEDAAYARDALAKAYSRFTWLVKIRNSL 382
 313 TSIDQVLEAFSEFTVEAKQEKVSTLTNVAQAYARDALAKAYSRFTWLVKIRNSL 372
 363 ASKDAESPSWSTVYGLDITGFEVPOHNSFEQPCINCYCEKIQOFLITLTKSEDEY 442
 373 KAQTIVRRK-----VMGLDITGFEVPOHNSFEQPCINCYCEKIQOFLITLTKSEDEY 427
 443 EAEGIAWEPVQYFNKIIICDLVEEFKGIISLDECELRPEATDLPFLKLEDTVPHP 502
 428 IREDIEWTHIDYFNNAIICDLIENNTNGLIAMDDECELRPEATDLPFLKLEDTVPHP 487
 503 HFLTH--KLADQKTRKSLDGEFRLIYAGVTVSVTGFLDKNDLFLRNLEKTMSSGM 560
 488 HFESMSKCSRFLLDTLPHSCFRIQHYAGKVLQVGEFVDKNNDLVRLDSLQMMWAGH 547
 561 PIMACPFKSELS--DKRREPVAPOFMSLLOVELIRSEKPAYICIRPNDAKQGRF 618
 548 ALISLPEEPGNPAVANKRPTASQKASATATLTKNPNYICIRPNDAKQGRF 607
 619 DEVLIRHGVKYLGLMELNVRRAQFAVRRKYEAFLOKYSCLPETHPMAGRPDQVAVL 678
 608 SESIVCHQIRVTLGLELVNVRRAQFAVRRKYEAFLOKYSCLPETHPMAGRPDQVAVL 667

679 VRLHGYPEEKYKMGKTKIFRFPPTLFEATEDSELEVRSQSLATKIOAMRGFMROKELRV 738
 668 FNELEIPEYEEKSFGSKIFIRNPRTLFQEDLRQRLDELATLQIKIRGKCKTHLLM 727
 739 KRSALCDSMMRGILGRKAKKAKKMAQOTIRLRIGF-----ILRHSPPCPNA---- 787
 728 KRSOVLAAWYRRYAQQRQYQOIKSSALVISTYIRGMARKILRELKHOKCKEAAVTIA 787
 788 ----- 787
 788 AVMHGTQARKRRRLKDEARKHAIIVIMAFVLSKARRELKLEKARRKHAVALTWAY 847
 788 -----FLDHYRASEFLMLRQLPR-NVLDTSWPPPALR 822
 848 WLGLKVRREYKFFRANAGKKIYFTLQRIQVYKYLEKKNMPSLIDKNWPSRPVFL 907
 823 EASELRELCKMKNW---KYCRSISPEMKQLOQKAVASIFPGKDNTPQSPRFLIS 879
 908 DSTH--KELKIFHLMKCKTRDOFTQOQKLYEKEASELEFKDKALPSSVGGPFG 965
 880 TRLGTEELSP---RVLSLQSEPIQYAVVPYKRYRGRYPRPQLLTPSAVVIYE--D 933
 966 AYL---EINKNPKYKKLKDAIEEKIIIAEYVKNINRANGSTSRIFLTNNLLADQKS 1022
 934 AKVQRIDYANLTSISVSLSDSLFVLAVQR-EDNKQGDVYLOSHVIEYTLK--TAL 989
 1023 GQIRSEVPLVDVTVKMSNDGFFAVHLKESGSAKSGDFLSSDLIMATKLYRTTL 1082
 990 SADRVN-NINI-----NQSITFAGPGPDGIIDFTSGSELLITAKNGHLAVVA 1038
 1083 SQTKQKNIETLSEDFLVQFQDKYCVKFIQGNQKNGSVP-----TCRKNRLEVA 1134

RESULT 13

S32404
 myosin heavy chain I, brain - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
 C:Accession: S32404; S33973
 R:Koslovsky, J.S.; Qian, C.; Jiang, X.; Mercer, J.A.
 FEBS Lett. 320, 121-124, 1993
 A:Title: Molecular cloning of a mouse myosin I expressed in brain.
 A:Reference number: S32404; MUID:93209370; PMID:8458427
 A:Accession: S32404
 A:Molecule type: mRNA
 A:Residues: 1-1094 <KOS>
 A:Cross-References: EMBL:X69987; NID:9312270; PIDN:CAA49604.1; PID:9166471
 R:Mercer, J.A.
 submitted to the EMBL Data Library, January 1993
 A:Reference number: S33973
 A:Accession: S33973
 A:Molecule type: mRNA
 A:Residues: 1-572, 'G', 574-762, 'G', 764-1094 <MER>
 A:Cross-References: EMBL:X69987
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: nucleotide binding; P-loop
 F:5-675/Domain: myosin motor domain homology <MOT>
 F:95-102/Region: nucleotide-binding motif A (P-loop)

Query Match 38.4%; Score 2075.5; DB 2; Length 1094;
 Best Local Similarity 41.6%; Pred. No. 1.1e-124;
 Matches 462; Conservative 183; Mismatches 347; Indels 119; Gaps 19;

27 VGVQDPVLENTSEAFIENLRRENLYITGVPVSVNPDQIYSRQHMERYR 86
 2 IGVGDVLEPL-NEETFDNLKRRFDHNEYITGYSVIVSNPYSRSLPIYSPEKVEDYR 60
 87 GVSFEVPEPLFAVDVTVRALTERRDOAVMISEGAGTEAKRRLQFAETCPAPE 146
 1083 SQTKQKNIETLSEDFLVQFQDKYCVKFIQGNQKNGSVP-----TCRKNRLEVA 1134
 147 RGVAVRDLRLSNPVLFAFNKATLRNDSSRFKYMVDQDFKAPVCGHILSYLTKS 206

Db 121 EYNQVKEOLQSNPVLAEAFNAKTYRNDNSSRFEGYMDIEDPKDPIGCVISNILEKS 180
 QY 207 RYVHONHGERNFHFVYOLLEGESEETLRGLERNPOSYLKVGQACAVSSINDKSDMK 266
 Db 181 RYVQKQPRGRNFHFVYOLLEGESEETLRGLERNPOSYLKVGQACAVSSINDKSDMK 239
 QY 267 VAKKALSVDFEDEDVLLSVASVHLGNTHFAADSDN---AQYTTENQKLYLRL 322
 Db 240 TVRNAMQVYGLDHEAEVLEVAALVGLKNEFPEESRVNGLDESKDKKNELEICEL 299
 QY 323 LGEVETTLREALTHRKIKGSELLSPNLDEQAARADALAKAVSRFTVWLVKINSIL 382
 Db 300 TSDIDVYLERASFPRVEAKQKYSTLTNVAQAYARADALANLXSRLESMVKNINSI 359
 QY 383 ASKDESPSRSTVYGLLDIYGFVFOHNSFOQCYNCNEKLOQLEFIETLKSEDEY 442
 Db 360 KATQVRRK-----VWGLDIYGFIEFEDNSFEQFLINCNKELQOIFIELTLKEOEY 414
 QY 443 EAEGLAMEPVQYFNKTIICDLYBEKFGIISILDECLRGATLTLETLEKEDYVKKP 502
 Db 415 IREDEWTHIDYFNNAITICDLENTNGILAMLEDECLRGATLTLETLEKEDYVKKP 474
 QY 503 HELTH--KLADQTRKSLDGEFRLHYAGEVYVYSGFLDKNNLLEFNLKEMCSKN 560
 Db 475 HESRMSKCSRFNLDTLPHSCFRIOHYAGVLYOEGFVDKNNLHYDLISQAMWAKGH 534
 QY 561 PMAQCFDKSELS--DKRPETVAOQKSLQVLEILSKREPATIRCTKPPDAQOPGRF 618
 Db 535 SLTKSLFEPGNPAKVLKRPPTRSSQFASVATLNMNITOTKNPNITRICKPDKKAHLE 594
 QY 619 DEVLIRHOYKYLGLMNLNVRAGAFYRRKYEAFLORYKSLCPETWPMAGRPQDQAVL 678
 Db 595 NESLVOHQIRYGLLENVAVRAGAFYRQAYEPCLEMYLCKQWPHMKGPARBSGEVL 654
 QY 679 VRHLGYPREYKMGRTKIFIRPKTLPATEDSLVVRQSLATKIQAMFGFIMROKFLY 738
 Db 655 FNELEIRVEHSEGRSKIFIRNPRITFOLEDLRKORLEDAVLQKTYGMCRCRHFILM 714
 QY 739 KRAIGIOSMWRGTLGRKRAAKRMAQOTIRLRIRF-----LRHSPREPENA---- 787
 Db 715 KRSQVIAAMRYTEOQKRIQOIKTSALVIOSTIRKMAKRLIRELHKOKREKAATTIA 774
 QY 788 -----FTLDH 792
 Db 775 AYWHGQARRELKRLKEARANDAVIWAUWMLGLKVRREYRKEFRANAGKKIYEPTLQR 834
 QY 793 VRASFLNLRLQRLP-NVLDTSWPTPRPLREASSELLRELCKNMW---KXCRSISPEW 848
 Db 835 IYQVYLLENKMKMPSIDKMPSPRYLELDSTH--KELKRLFIHLMRCKKTRDOFTDQ 892
 QY 849 KOOLQOKAVASEIFKCKNDYPOSVPLRIFSTRIGTEESIP-----RVLOSLSSEPLOYAV 904
 Db 893 KLIYEKLESELEFKKALYPSVQOPQAVL---EINKNPKYKKLDAIEKIIIAE 949
 QY 905 PVAKYDRKGYRPRROLLTPSAVIVE--DAKQORIDYANLTGISVSSLSLFLVHY 962
 Db 950 VVKIRANRANKSTRIFLTNNLLADQSGQIKSEVPLVDYTVKSSQNDGFAVHL 1009
 QY 963 QR-EDMKQGDVLODHYETLT---TALSADRN-NINI-----NQSITFA 1007
 Db 1010 KESSEASAKGDFLFSSDHLIEMATKLYRTLLSQTKKLNIEISDFLVOPRODKVCVKFI 1069
 QY 1008 GGGRDGIIDFTSGSELLITKAKNGHLAVYA 1038
 Db 1070 QGNQKNQSV-----TCKRKNRLLEVA 1092

RESULT 14

A33620
 myosin heavy chain I, brush border - chicken (fragment)
 N:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Gallus gallus (chicken)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002

C:Accession: A33620
 R:Garcia, A.; Coudrier, E.; Carboni, J.; Anderson, J.; Vandekerckhove, J.; Mooseker, M.
 J. Cell Biol. 109, 2895-2903, 1989
 A:Title: Partial deduced sequence of the 110-kD-calmodulin complex of the avian inces
 A:Reference number: A33620; MUID:90078325; PMID:2687288
 A:Accession: A33620
 A:Molecule type: mRNA
 A:Residues: 1-1000 <GAR>
 A:Cross-references: GB:X58479; NID:g63075; PID:CA41388.1; PID:g63076
 C:Superfamily: brush border myosin heavy chain I; myosin motor domain homology
 C:Keywords: actin binding; hydrolase; intestine; nucleotide binding; P-loop
 F:1-639/Domain: myosin motor domain homology <MMO>
 F:59-66/Region: nucleotide-binding motif A (P-loop)
 F:530-551/Region: actin binding #status predicted
 F:681-1000/Domain: carboxyl-terminal <CTD>
 F:65/Binding site: ATP (uys) #status predicted

Query Match 35.3%; Score 1904; DB 1; Length 1000;

Best Local Similarity 41.6%; Pred. No. 9.2e-114;

Matches 415; Conservative 191; Mismatches 358; Indels 34; Gaps 16;

QY 64 VLVSYNPRDQIQRSMERKRYGVSYEPHFLFADAVYRALRERRDOAVMISGES 123
 Db 2 VVISVNPYKPLPIYTPKEVEYHNCNFAVPHYALADATKSLRDRDQCLITGES 61
 QY 124 GAGKTEATKRLQFYAETCPAPERGAVERDLQSNPVLAEAFNAKTYRNDNSSRFEGKYM 183
 Db 62 GAGKTEASKLYMSVYAAVSSKGEVDKKBGLQSNPVLAEAFNAKTYRNDNSSRFEGKYM 121
 QY 184 DVQDFKCAPVGHILSTYLEKSRVYHONGERNFHFVYOLLEGESEETLRGLLENPQ 243
 Db 122 DVEEPFKDPPGAGVISNLTLEKSRVYHONGERNFHFVYOLLEGESEETLRGLLENPQ 181
 QY 244 SYLVLYGQCAKSSINDKSDMKVRRKALSYIDTEDEDVLLSVASVHLGNI----H 299
 Db 182 HYGL-NEKSVLYGMDAANFRAMQDAMALIGRAPAVTALLEVAVLKGIVKSSS 240
 QY 300 FADEDSNAQVTTENQKLYLRLGVESTYREALTHRKIIAKGEBELSPNLDEQAAYAR 359
 Db 241 FOASGMEASSIAEPRELOEISQILGIDPSTEOALCSRTVAVRDESVLTALYSQGYGR 300
 QY 360 DALAKAVSRFTVWLVKINSLSKQAESVSWSTVYGLLDIYGFVFOHNSFEQFCI 419
 Db 301 DALAKNIYSRLFEDVWVNRINTSIOVKPg-----KQKVMGLDIYGFIEFODNGFEPFI 355
 QY 420 NYCNEKIQQLFTELTLKSEDEYEAEIGAMEPVQYFNKTIICDLYBEKFGIISILDEEC 479
 Db 356 NYCNEKIQQLFTELTLKSEDEYEAEIGAMEPVQYFNKTIICDLYBEKFGIISILDEEC 415
 QY 480 LRPGATDLPLEKLEDTYVQPHFLYHKLADQK--TRKSLDRGEFRLHYAGEVYVYSGFL 537
 Db 416 LRPGVNDPTFTTKLNQIFASHKRYSEKETINAHVYDVSLPLRCFRINHAGVYTVNT 475
 QY 538 GLDKNNLDPNFKLETKSSNPIMAQCFDKS--ELSDKRPETVAOQKSLQVLEILSKREP 595
 Db 476 GETERNNDLRLDSQAMAAARHILBSLFEGBDQPRQSLKLPPTTOSQAFASVATLMKN 535
 QY 536 LSKSEPAYIRCIKPPDANQOPGREDEVILRHQVYKYLGLMNLNVRAGAFYRRKYEAFLOR 655
 Db 536 LYSKNPNYIRCIKPPDNTKAMLEFPDILVAQVRYGLIMENVRVRAAGAFYRQLOPFLER 595
 QY 656 YKSLCPETWPMAGRPQGVAVVYHGLGKDEEYKMGRTKIFIRPKTLPATEDSLVVR 715
 Db 596 YKMLSRKTWPRWTGDRGAELVLELKEFPEELAYGHTKIFISAPFLFLEKRRQORV 655
 QY 716 QSLATKIOAAMRGFMRQFLKVRSAICIOSMWRGTLGRKRAAKRMAQOTIRLRIRF 775
 Db 656 AELATLQKMFQWCCCKRYQLMKRSQILLISAMFGHQRRYRQMKSVYLLQAYARGW 715
 QY 776 IIR-----HSRQPE-NAEFLDVHVASFLINLRQL-PRNVLDTSWPTPR-PALKEAS 825
 Db 716 KTRRMVRRYFRSDACTRLSNFYRRMVOYKIMGLQKNLPMAVILDRTPAPAYKFLSDAN 775

Qy 826 ELRLRLCKNMWYKRSISPEWKOOLQAKAVASEIFGKKNDYPOSVPRLLFSTRIG-T 884
Db 776 QELKSIYFRMKCKKREQLTPQORAMLOAKLMPRLQFDRKALALQSQPFREYIGLT 835
Qy 865 EELSPRVQLSGSEPIQYAVPVYKDKRGYKPRRQLLTSPSAVYVED--AKYKORIDY 942
Db 836 QNRKQKLOAVAKKLVMAEAVOKYVRANGKTVPRLLLTTEHVLADPKAAQPKMYLSTL 895
Qy 943 ANLGISVSSLSDSLFLVHQREDNK-QKGDVVLQSDHVIETLRK---TALSAD-RVNNI 997
Db 896 CDIQASVSRPSDGLALHLKETSTAGKGLLVSPHLIELVTRLHOTLMDAVAOALPL 955
Qy 998 NI-NGSITFAGGPGRDGIIDFT--SGSELLITKAKNGH 1033
Db 956 SIADQFTRFPKGDVAAYVVEESAKGGGVPCYKRGSH 993

RESULT 15
A: myosin I myr 4 - rat
A: Accession: A53933
C: Species: Rattus norvegicus (Norway rat)
C: Date: 07-Oct-1994 #sequence_Revision 07-Oct-1994 #text_Change 02-Feb-2001
C: Accession: A53933
R: Baehner, M.; Kroschewski, R.; Stoeckli, H.E.; Behrmann, T.
J. Cell Biol. 126, 375-389, 1994
A: Title: Rat myr 4 defines a novel subclass of myosin I: identification, distribution,
A: Reference number: A53933; MUID: 94308268; PMID: 8034741
A: Accession: A53933
A: Status: Preliminary
A: Molecule type: mRNA
A: Residues: 1-1006 <BAE>
C: Cross-references: GB:X71997; NID:9516395; PID:9516396
C: Superfamily: brush border myosin heavy chain I; myosin motor domain homology
C: Keywords: nucleotide binding; P-loop
F: 12-682/Domain: myosin motor domain homology <AMOT>
F: 102-109/Region: nucleotide-binding motif A (P-loop)

Query Match 34.5%; Score 1861; DB 2; Length 1006;
Best Local Similarity 41.7%; Pred. No. 5,3e-111;
Matches 426; Conservative 158; Mismatches 301; Indels 136; Gaps 26;

Qy 28 GVQDFVLLNFTSEAFLENNRRRENLIYTYGVPVLSVNPRLDIYSRQHMERYG 87
Db 10 GKADVLMDT--VSMDEFMANLRLREKGRITFTIGEVVSVNPYKVLNITGRDITQYKG 68
Qy 88 VSFYVPPHPLAVADVTRALRTERDQAVMISGESGAKTEATKRLQFYAETCPAPER 147
Db 69 RELYERPHPLAIDAAVKAMKRRSKDTCMISGESGAKTEASKTYIMQYIAATNPQR 128
Qy 148 G--GAVRDLLOSNPVLAFAFGNAKTLRNDSSRFQKYMDOVDFKAPVGGHLSYLEK 205
Db 129 AELENVKMLKSKCVLAFAFGNAKTNRNDSSRFQKYMINDFDRGDPIGHINNYLLEK 188
Qy 206 SRVHQNHGERNFHVYQLLEGSEETLRRLGLERNPQSYLYVKGQCAKVSINDKSDW 265
Db 189 SRVIVQOGERSFHSFYQLLOGSROMLSHLQKSSVNYIRVGAQLK--STINDAEF 247
Qy 266 KYMKALSYIDFTEDEVEDLSIVASVHLGNINPADEDS----NAQVTENQIKYLTR 321
Db 248 KVVADAMKVIGFKPREIYVKKILAILHNLKFIYDGPPLIENGKRV---VSYTAE 302
Qy 322 LIGVEGTLREALTRKIIAKGEELSPLELQ--AAVADALAKAVSRFTWLVKRLNR 380
Db 303 LLSTADNVEKALLR--IVATGRDITIDKHTEQESYGRDAFAKAIYERLECWITRIND 361
Qy 381 SLASHDAESPWRSTVVLGLDIYGFVEVQHNSEFQPCINCNKIQQLFTELTKSBOE 440
Db 362 IIEVKNYDVTYVHGKTVIGVLIYGFELFDNNSFEQFCINCNKIQQLFTELTKSBOE 421
Qy 441 EYEAGIAMEPVQYVNNKIIICDLVEKFGIISIDECILRGEXTDLTFLEKLEDTYKP 500
Db 422 EYQREGIPMKHIDYFNNOIYDLVBOHQHGIILDLDDACMNVGKVTGDMFLDALNSKLGK 481

Qy 501 HPFHLTKLADQKTRKSLDRGEERLLHYAGEVTVSYTGFLDKNNDLLFRNLKETMCSNN 560
Db 482 HGHFSSKKTICASDKILEFDR--DFRINHYAGDVYVYVIGFDKKKOTLLFQDFKRLMYNSN 540
Qy 561 PMAQCFDKSELS---DKRPETVAQFKMSLQVLEILSKSPAYIRCIKPNDAKQGR 617
Db 541 PVLKMPPEKRLSITEVTKRPLTAATLFFKSMIALVDNLASKPEYVYRCIKPNDKSPOI 600
Qy 618 FDEVLTRHQVYGLMENNLPVRAGAVRRKYKAFIQRKSLCPETWPMAGROPQGVAV 677
Db 601 FDEGRGHQVEYIGLENNVRRAGFAFROTYSKPLRRKMYSEFTWPN--HDLPSDKEAV 659
Qy 678 --LVRLHGYRPEEYKMGRTKIFTRPKTLFATEDSLEVRROSLATKIQAMRGFHWROKF 735
Db 660 KKLIERGFG--DVAAKGTIFTRTRITLPLE---ELRQMLV-----699
Qy 736 LRVKSAICIQSMWRGTLGRKRAKRRKMAQITRLIRGFIHNSRCPENAEFLDHVRA 795
Db 700 ---RVVLFQKVMRGTLARMR--YKRTKALITIRYRRKYK-----737
Qy 796 SFLNLRLRL--PRNYLD---TSWPTPPALREASELLRELCKNMWYKRSISPEWK 849
Db 738 SYLHEVARRHGYKNNMDYGGKHYKWPYPRVLRFEEDALQSTFNRBRASOLITIPASDI 797
Qy 850 QOLQOKAVASEIFGKKNDYPOSVPRLLFSTRIGTEBISPRVQLSGSEPIQYAVPV-- 907
Db 798 PQYRAKVAAMEMLKQGRADL--GLQRAWENYILASKPDP--QTSGT-----FVYVANE 847
Qy 908 ----KT-----DKRGYKPRRQLLTSPSAVYVEDAKY 936
Db 848 LKRKDKYMNVLFSCHVARKVNRFSKVEDRALFVTDRLHYKMDP-----TRQYKY 895
Qy 937 KORIDYANLGISVSSLSDSLFLVHQREDNKQGDVVLQSDHVIETLRKALSADRVNN 996
Db 896 MKTIPLYNLGLSVSNKQDLVYFHT--KDNK-----DLIVCLFSQPTHESTRIGE 944
Qy 997 I 997
Db 945 L 945

Search completed: July 7, 2003, 14:24:54
Job time: 59.1849 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 29.5472 seconds
(without alignments)
1465.497 Million cell updates/sec

Title: US-09-893-371-1
Perfect score: 5398
Sequence: 1 MRYRASALGSDGVRTMESA.....LITKAKNGHLAVVAPRLNSR 1044

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SWISSProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5320	98.6	1028	MYIC_MOUSE	Q9W117 mus musculu
2	5143	95.3	1028	MYIC_HUMAN	Q00159 homo sapien
3	2626	48.6	1026	MYIB_DROME	Q23979 drosophila
4	2106.5	39.0	1043	MYHL_BOVIN	P10568 bos taurus
5	2090.5	38.7	1107	MYIA_MOUSE	P46735 mus musculu
6	2080	38.5	1136	MYIA_RAT	Q05096 rattus norv
7	2064.5	38.2	1043	MYHL_HUMAN	Q9UDC5 homo sapien
8	2005.5	37.2	1045	MYHL_CHICK	P47807 gallus gall
9	1987	36.8	909	MYHL_MOUSE	Q88829 mus musculu
10	1932.5	35.8	842	MYHL_RAT	Q62374 rattus norv
11	1800.5	33.4	1011	MYIA_DROME	Q23978 drosophila
12	1739	32.2	1003	MYSE_MOUSE	Q03479 dictyostell
13	1726.5	32.0	1099	MYFE_MOUSE	P70248 mus musculu
14	1692.5	31.4	1111	MYSD_DROME	P34092 dictyostell
15	1601	29.7	1113	MYSD_DROME	P34109 dictyostell
16	1588.5	29.4	994	MYSA_DROME	P22467 dictyostell
17	1574.5	29.2	1147	MYSB_ACACA	P19706 acanthamoeb
18	1523.5	28.2	1168	MYSC_ACACA	Q04439 saccharomyc
19	1501	27.8	1219	YMG9_YEAST	P42522 dictyostell
20	1453	26.9	1181	MYSC_YEAST	P36006 dictyostell
21	1427.5	26.4	1273	MYSA_HUMAN	Q13402 homo sapien
22	1354	25.1	2215	MYA_BOVIN	P79114 bos taurus
23	1325	24.5	2052	MYIO_HUMAN	P97479 mus musculu
24	1284.5	23.8	2058	MYIO_HUMAN	Q9ND67 homo sapien
25	1279	23.7	2116	MYSC_DROME	P08799 dictyostell
26	1263	23.4	1742	MYSC_HUMAN	Q9Q424 homo sapien
27	1259	23.3	3511	MYIS_MOUSE	Q9Q424 mus musculu
28	1243.5	22.0	3530	MYIS_HUMAN	Q9Q424 mus musculu
29	1233.5	22.0	1509	MYSN_ACACA	P05659 acanthamoeb
30	1225.5	22.7	1853	MYSA_MOUSE	Q99104 mus musculu
31	1224.5	22.7	1828	MYSA_RAT	O99YF3 rattus norv
32	1224.5	22.7	2245	MYSD_DROME	P54697 dictyostell
33	1224.5	22.7	2245	MYSD_DROME	P54697 dictyostell

34	1217.5	22.6	1938	MYA_BOVIN	P24733 aequipecten
35	1217	22.5	1855	MYSA_HUMAN	Q9Y411 homo sapien
36	1216.5	22.5	1846	MYSB_RAT	P70569 rattus norv
37	1206	22.3	1829	MYIB_CHICK	Q02440 gallus gall
38	1206	22.3	1978	MYHB_CHICK	P10587 gallus gall
39	1202	22.3	1935	MYH7_RAT	P79293 sus scrofa
40	1198	22.2	1935	MYH7_RAT	P02564 rattus norv
41	1192	22.1	1972	MYHB_HUMAN	P35749 homo sapien
42	1192	22.1	1939	MYSN_DROME	Q99323 drosophila
43	1190	22.1	1939	MYHB_HUMAN	P13533 homo sapien
44	1190	22.0	1972	MYHB_RAT	P35748 oryctolagus
45	1187.5	22.0	1976	MYHA_BOVIN	Q27991 bos taurus

ALIGNMENTS

RESULT 1
MYIC_MOUSE
ID MYIC_MOUSE STANDARD: PRT: 1028 AA.
AC Q9W117; 008834; 008571; Q9QW54;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin Ic (Myosin I beta) (MIMD).
GN MYIC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC STRAIN=BA/2J; TISSUE=Brain;
RX MEDLINE=97325785; PubMed=9182797;
RA Hamilton B.A., Smith D.J., Mueller K.L., Kerrebrock A.W.,
RA Bronson R.T., van Berkel V., Daly M.J., Kruglyak L., Reeve M.P.,
RA Nemhauser J.L., Hawkins T.L., Rubin E.M., Lander E.S.,
RT "The vibrator mutation causes neurodegeneration via reduced expression
of P1TP alpha: positional complementation cloning and extragenic
suppression."
RT Neuron 18:711-722(1997).
RL [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawata J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Buit C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kanliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima R., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 408:685-690(2001).
RN [3]
RP SEQUENCE OF 1-807 FROM N.A. (ISOFORM B).
RC STRAIN=BA/2J; TISSUE=Cochlea;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., Amraoui A.E., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hamel C., Fitzames C., Levi-Acobas F., Depetris D., Mattei M.-G.,
RA Weil D., Pujol R., Petit C.,
RT "Cloning of the genes encoding two murine and human cochlear

RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 RN [4]
 RP SEQUENCE OF 146-593 FROM N.A.
 RX MEDLINE-93194946; PubMed-8449986;
 RA Sherr E.H., Joyce M.P., Greene L.A.;
 RT "Mammalian myosin I alpha, I beta, and I gamma: new widely expressed
 RL genes of the myosin I family.";
 J. Cell Biol. 120:1405-1416(1993).
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
 CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
 CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
 CC SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IO DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U96723; AAC53264.1; -;
 CC EMBL: AK004743; BAB23524.1; -;
 CC EMBL: U96726; AAC60758.1; -;
 CC EMBL: X99638; CA67956.1; -;
 CC HSSP: P08799; 1MDN.
 CC MGD: MGI:106612; Myo1c.
 CC InterPro: IPR000048; IO_region.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00063; myosin_head; 1.
 CC Pfam: PF00612; IO; 3.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART: SM00015; IO; 2.
 CC SMART: SM00242; MYSC; 1.
 CC PROSITE: PS50096; IO; 2.
 CC Myosin: ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
 CC MultiGene family; Alternative splicing.
 CC FT DOMAIN 1 683
 CC FT DOMAIN 699 722
 CC FT DOMAIN 723 751
 CC FT NF_BIND 105 112
 CC FT VARSPLIC 897 1028
 CC
 CC RR -> GG (IN REF. 3).
 CC SKO -> RRR (IN REF. 3).
 CC T -> A (IN REF. 4).
 CC C -> F (IN REF. 4).
 CC VM -> LL (IN REF. 4).
 CC R -> A (IN REF. 4).
 CC LAS -> VPA (IN REF. 4).
 CC Q -> R (IN REF. 4).
 CC VKP -> IKR (IN REF. 4).
 CC M -> T (IN REF. 4).
 CC S -> G (IN REF. 4).
 CC O -> R (IN REF. 3).
 CC R -> G (IN REF. 3).
 CC E -> D (IN REF. 3).
 CC
 CC 1028 AA; 118155 MW; 2CF5535D816884DF CRC64;
 CC
 CC Query Match 98.6%; Score 5320; DB 1; Length 1028;
 CC Best Local Similarity 100.0%; Pred. No. 0;
 CC Matches 1028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 MESALTARDRYGVDFVLENTSEAFIENLRFRREFNTIYTGIVLVSVNPDLOI 75
 DB 1 MESALTARDRYGVDFVLENTSEAFIENLRFRREFNTIYTGIVLVSVNPDLOI 60
 QY 77 YSRQHMERKRGVSFEVPPHFAVADIVYRALTERDDQAVMISGESGACKTAIRKLQ 136
 DB 61 YSRQHMERKRGVSFEVPPHFAVADIVYRALTERDDQAVMISGESGACKTAIRKLQ 120
 QY 137 FYAETCAPENGAVRRLLOSNPVLEAFNAKTLRDNDSRFGKVDYDFDKGAPVGG 196
 DB 121 FYAETCAPENGAVRRLLOSNPVLEAFNAKTLRDNDSRFGKVDYDFDKGAPVGG 180
 QY 197 HILSYLEKSRVYQNHGERNFHFVQLLEGEEETLRGLERNPOSYLYVKGQCAKY 256
 DB 181 HILSYLEKSRVYQNHGERNFHFVQLLEGEEETLRGLERNPOSYLYVKGQCAKY 240
 QY 257 SSINDKSDMKVKKALSYIDTDEVEDLSIVASVHLGNHRAEDSNQAVTENOL 316
 DB 241 SSINDKSDMKVKKALSYIDTDEVEDLSIVASVHLGNHRAEDSNQAVTENOL 300
 QY 317 KYLRRLGVEGTTLRALTHRKITAKGELLSPNLBOAAVARDALAKAVYSFTWLV 376
 DB 301 KYLRRLGVEGTTLRALTHRKITAKGELLSPNLBOAAVARDALAKAVYSFTWLV 360
 QY 377 KINSLASKDAESPSWSTVGLDIDYGFEEVFOHNSFEQFCINCEKLUQLEFELTK 436
 DB 361 KINSLASKDAESPSWSTVGLDIDYGFEEVFOHNSFEQFCINCEKLUQLEFELTK 420
 QY 437 SEQEEYEAEGTAMPEVOYFNKKIICDIVEEFKGIISLIDECRPREADTFLFLEKD 496
 DB 421 SEQEEYEAEGTAMPEVOYFNKKIICDIVEEFKGIISLIDECRPREADTFLFLEKD 480
 QY 497 TYKPRPHFLTKRLADOKTRKSLDRGEFRLIYAGEVTVYTGFLDKNNDLFRNLKETWC 556
 DB 481 TYKPRPHFLTKRLADOKTRKSLDRGEFRLIYAGEVTVYTGFLDKNNDLFRNLKETWC 540
 QY 557 SSANPIMAOCPDKSLSKKRPETVAVQPKFSILQVLEILRSKPATYRCIKPDADKOPG 616
 DB 541 SSANPIMAOCPDKSLSKKRPETVAVQPKFSILQVLEILRSKPATYRCIKPDADKOPG 600
 QY 617 RFDEVLIHROVKYLLGIMENLRVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPDQVA 676
 DB 601 RFDEVLIHROVKYLLGIMENLRVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPDQVA 660
 QY 677 VLVRLHGYKPEEYKMGRTKIFIRPKTLFATEDSLVAVROSLATKIQAAWGFHWROKFL 736
 DB 661 VLVRLHGYKPEEYKMGRTKIFIRPKTLFATEDSLVAVROSLATKIQAAWGFHWROKFL 720
 QY 737 RYKRSALICQSWMRGTLGRKAARKKMAAOTIRLRIRGFIILRHSRPCENAFFLDHVRAS 796
 DB 721 RYKRSALICQSWMRGTLGRKAARKKMAAOTIRLRIRGFIILRHSRPCENAFFLDHVRAS 780
 QY 797 FLINLRROLPRNVLDTSWTPPALREASSELLRELCKMNMWYKCRSISPWKQOLOOKA 856
 DB 781 FLINLRROLPRNVLDTSWTPPALREASSELLRELCKMNMWYKCRSISPWKQOLOOKA 840
 QY 857 VASEIFKGGKNDYVPRPLISTRLGTEBETSPRVLQSLSEPPLOYANPVYKVKRKGKP 916
 DB 841 VASEIFKGGKNDYVPRPLISTRLGTEBETSPRVLQSLSEPPLOYANPVYKVKRKGKP 900
 QY 917 RPROLLTPPSAVIYEDAKVKORIDYANITGIVSLSLDFVLVHVGREDKOKGDVYLQ 976
 DB 901 RPROLLTPPSAVIYEDAKVKORIDYANITGIVSLSLDFVLVHVGREDKOKGDVYLQ 960
 QY 977 SDHVLETLTKTALSADRVNNININGSTIFPAGPGRDIIDFTSGSELLITKANKGLAV 1036
 DB 961 SDHVLETLTKTALSADRVNNININGSTIFPAGPGRDIIDFTSGSELLITKANKGLAV 1020
 QY 1037 VAPRLNSR 1044
 DB 1021 VAPRLNSR 1028

```

RESULT 2
MYIC_HUMAN STANDARD: PRT; 1028 AA.
AC 000159:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin IC (Myosin I beta) (MMI-beta) (MMIB).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., Amroult A.E., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hamel C., Fitzames C., Levi-Acobas F., Depetris D., Mattel M.-G.,
RA Weil D., Pujol R., Petit C.;
RT "Cloning of the genes encoding two murine and human cochlear
RT unconventional type I myosins.";
RL Genomics 40:332-341(1997).
CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X8507; CAA67131.1; -.
DR HSSP: P08799; IMND.
DR GeneW: HGNC:7596; MYO1B.
DR Genew: HGNC:7597; MYO1C.
DR MIM: 606538; -.
DR
DR InterPro: IPR000048; IQ_region.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00612; IQ; 3.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 2.
DR SMART: SM00242; MISC; 1.
DR PROSITE: PS50096; IQ; 2.
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
KW Multigene family.
FT DOMAIN 1 683 MYOSIN HEAD-LIKE.
FT DOMAIN 699 722 IQ 1.
FT DOMAIN 723 751 IQ 2.
FT NP_BIND 105 112 ATP (POTENTIAL).
SQ SEQUENCE 1028 AA; 118037 MW; 0E9C3680527F85C6 CRC64;
Query Match 95.3%; Score 5143; DB 1; Length 1028;
Best Local Similarity 96.2%; Pred. No. 5.6e-319;
Matches 989; Conservative 23; Mismatches 16; Indels 0; Gaps 0;

```

```

OY 17 MESALTARDVGVDFVLENTSEAFIENLRFRFENLITYTYIGPVYVNPYRDLQI 76
DB 1 MDSALTARDVGVDFVLENTSEAFIENLRFRFENLITYTYIGPVYVNPYRDLQI 60
OY 77 YSRQHEERYRGVSFEVPEPHLEAVADYTYRALTERRRQAVMISGESGAGTEATKRLQ 136
DB 61 YSRQHEERYRGVSFEVPEPHLEAVADYTYRALTERRRQAVMISGESGAGTEATKRLQ 120

```

```

OY 137 FYAETCPADRGCAVRDLQSNPVLAEFGNAKTLRDNSSRFGRKYMVQDFDKGAPYGG 196
DB 121 FYAETCPADRGCAVRDLQSNPVLAEFGNAKTLRDNSSRFGRKYMVQDFDKGAPYGG 180
OY 197 HILSYLLEKSRVYVHNGHGRNHFVYQLLEGESEETLRGLERNPQSYLYLVKQCAKV 256
DB 181 HILSYLLEKSRVYVHNGHGRNHFVYQLLEGESEETLRGLERNPQSYLYLVKQCAKV 240
OY 257 SSINDSPMKVYKRAKLSVIDFTEDEVEDLSTIVASVLIHGNHFAADSDNAQVYTEQL 316
DB 241 SSINDSPMKVYKRAKLSVIDFTEDEVEDLSTIVASVLIHGNHFAADSDNAQVYTEQL 300
OY 317 KULTRLVGEVGTTLREALTHRKTIKAGBELLSPNLEQAAVARDALAAVYSRFTWLV 376
DB 301 KULTRLVGEVGTTLREALTHRKTIKAGBELLSPNLEQAAVARDALAAVYSRFTWLV 360
OY 377 KINRSIASKDAESPMSRSTVYGLLDIYGFVFQNSPEQPCINCKEKLQDLFTELTK 436
DB 361 KINRSIASKDAESPMSRSTVYGLLDIYGFVFQNSPEQPCINCKEKLQDLFTELTK 420
OY 437 SEQEEYEAEGIAEPEVQYFNKKIICDLVEEKFGLIISLDECELRGEATDLTELEKED 496
DB 421 SEQEEYEAEGIAEPEVQYFNKKIICDLVEEKFGLIISLDECELRGEATDLTELEKED 480
OY 497 TYKPHPHLTAKLDOKTRKSLDGEFFLLHYAGVYTSYGFIDKNNDLFRULKETMC 556
DB 481 TYKPHPHLTAKLDOKTRKSLDGEFFLLHYAGVYTSYGFIDKNNDLFRULKETMC 540
OY 557 SSMPDMAOCEPKSELSDKRRPETAQFQKSLDLVLEILSKPEPAYIRCIKPDADKOPG 616
DB 541 SSMPDMAOCEPKSELSDKRRPETAQFQKSLDLVLEILSKPEPAYIRCIKPDADKOPG 600
OY 617 RFEDEVLIHQVYKYLGLMENLVRNAGFAVRRKYAFLQRYKSLCPETWMAAGRPQDVA 676
DB 601 RFEDEVLIHQVYKYLGLMENLVRNAGFAVRRKYAFLQRYKSLCPETWMAAGRPQDVA 660
OY 677 VLVRHLGYKPEEYKMGRTKIRPPTLFAFEDSLVYRROSLAKTQAMNGEHWKQFL 736
DB 661 VLVRHLGYKPEEYKMGRTKIRPPTLFAFEDSLVYRROSLAKTQAMNGEHWKQFL 720
OY 737 RVKRSALICQSWMNGTIGRRRAAKRMAAQIRLIRGFIIRHSRCPENAFELDHVRAS 796
DB 721 RVKRSALICQSWMNGTIGRRRAAKRMAAQIRLIRGFIIRHSRCPENAFELDHVRAS 780
OY 797 FLNLRLQPLNNVLDTSMPRPALREASSELRLCLCKNMWKCISPEMKOQLQOKA 856
DB 781 FLNLRLQPLNNVLDTSMPRPALREASSELRLCLCKNMWKCISPEMKOQLQOKA 840
OY 857 VASEIFGAKDNDYQSVPRLEISTRLGTEISPRVLSLSEPIQYAVPVYKDYDRGKYP 916
DB 841 VASEIFGAKDNDYQSVPRLEISTRLGTEISPRVLSLSEPIQYAVPVYKDYDRGKYP 900
OY 917 RPRLLTTPSAVYVEDEAKYKQRIQIDYANLIGISVSSLSDSLFLVHVRQEDNKGQDVIQ 976
DB 901 RSROLTLTPNAVYVEDEAKYKQRIQIDYANLIGISVSSLSDSLFLVHVRQEDNKGQDVIQ 960
OY 977 SDHYIETLTALTSADRVNININOGSITFAGGGRGDIIDTSGSLLTTKKNGLAV 1036
DB 961 SDHYIETLTALTSADRVNININOGSITFAGGGRGDIIDTSGSLLTTKKNGLAV 1020
OY 1037 VAPRLNSR 1044
DB 1021 VAPRLNSR 1028

```

```

RESULT 3
MYIB_DROME STANDARD: PRT; 1026 AA.
AC Q23979: Q9W0H0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Myosin IB (MIB) (Brush border myosin IB) (BBMIB).

```

05 MYO61F OR CG9155.
 06 Drosophila melanogaster (Fruit fly).
 07 OS Eukaryota; Metazoa; Arthropoda; Mendibulata; Pancrustacea; Hexapoda;
 08 Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 09 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 10 NCBI_TaxID=7227;
 11 [1]
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 51

[illegible]

```

Db      121 FIAAGSGNQTVEGVKDKLLKSNPVLAEFGNAKTNNDNSRRGKYMIDQDFKGAIPGG 180
Qy      197 HILSTLLEKSRVYVHNGHGERNFVVFQLEGGEEFLRLGLERNPQSTLYLVKGCQAVY 256
      181 NITLVLEKSRVYVHNGHGERNFVVFQLEGGEEFLRLGLERNPQSTLYLVKGCQAVY 240
Qy      257 SSINDKDMVKMKALSVIDEFTEDELLSTVASHLGNHFEADEGSMNQVYTNEL 316
      241 TRINDADSFQVQALTVIDEFTEDELLSTVASHLGNHFEADEGSMNQVYTNEL 299
Qy      317 KYTRLGVEGTTLRALTRHRTIAKGEELLSPLEAOAYADALAKAVSRFTWLR 376
      300 VTAARLLGVASLELAALTRHRTIDARGDVATSPNGELAIYARDAALAKAVSRFTWLR 359
Qy      377 KIRSLASDAESPSRSTTVGLDLYGFEPVPHNSFQFCINYNENKLOQLEFLTLK 436
      360 RLIVSLQAKETRA---SRNNVGLDLYGFEPVPHNSFQFCINYNENKLOQLEFLTLK 416
Qy      437 SEDEYRARGIEMIPVEYFNKYLICLIEKKHGIISIDECLRGEPDTDTFLEKLTQ 496
      417 SEDEYRARGIEMIPVEYFNKYLICLIEKKHGIISIDECLRGEPDTDTFLEKLTQ 476
Qy      497 TYKPRPHFLTKLAQDKTRKSLDRGFRLHLAGVYTVSGFLDKNDLFRNLKETWC 556
      477 KLAQHNYVCHERAPAHIKKIMLDFEFLVHYAGEVTVSGFLDKNDLFRNLKETLS 536
Qy      557 SSMNPIMAOCPDSEKSDKKRPETVATOKMSLDLVELLRKEPATYICIRPNDAKOG 616
      537 KAGNGIVRSCFPKEKLSRKREPTATQORASINNMIDLMCKESBYICIRPNDAKOG 596
Qy      617 RDEVLIRHGVYKYLGLMENLRYRAGFAVRRKYEAFQYKSLCEPTWPMW---AGRPDQ 674
      597 VENDELVLHGVYKYLGLMENLRYRAGFAVRRKYEAFQYKSLCEPTWPMW---AGRPDQ 656
Qy      675 VAVLVNHLGKYEKYEKMGRTKIFIRPKTLFTEDESLVYRQSLATKIQANRGFHWK 734
      657 VOQLVADLGMDEKRYVGETKLEIRPRLTFDEYAYOKKEHIAIADHKKGMORRK 716
Qy      735 FLVKKSAICISQSMWGTGLRRKAKRKAQAOTIRRLINGFLIRSPRCEPAFLDHR 794
      717 YLKLRAOVILMOSYCRKLAQAQAKRRRAKIRAFIKGFTTRDAPNGEENFTANK 776
Qy      795 ASFLNLRQDPRNVLDTSWPTPPALRASFLRELCKNMWYKRSISPEWQALQ 854
      777 RMMLRLAKELPTKYLDSMPHAPGCEASGILRHHLHRLIARLYRLKLPQQRQEL 836
Qy      855 KAVASEIFKGGKNDYPOSYRPLFISRLCTEET---SPRYLOSLSSEPQYAVPVKTD 911
      837 KYLAERVFEGKKNNTASSVSTWFOEDRIPKEHIOHVNDVASTFSEDLKYQSFCTKDR 896
Qy      912 KGKPRPRLTLTPSAVNVLEDKAKQ-----RIDYANLIGISVSSLSDSFLVHOR 964
      897 HGKSRDRFLLSNNAIYLDGKTYKOKHRLPLDKIDF-----TLTNHNDLWIRPL 950
Qy      965 ENNKOKGVYLOSNAVETLTALSDRVNNINI--NOSGITFAGPGRDGIIDTSGSE 1023
      951 DUKKRGDILILIPRIESTYIIDLTVGASIVSDRNSLEHNVKKGVIDIQTGAE 1010
Qy      1024 LITRAKNGHLAVY 1037
      1011 PGVYDK-GHLVIT 1023

```

```

RESULT 4
MYHL_BOVIN STANDARD: PRT: 1043 AA.
AC P10568: Q9NS38: 01-JUL-1989 (rel. 11, Created)
DT 01-JUL-1989 (rel. 11, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Brush border myosin I (BBMI) (Myosin I heavy chain) (MHC)
EN (Brush border 110-kDa protein).
GN MYHL.

```

```

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8033016; PubMed=3667594;
RA Hoshimaru M., Nakajishi S.;
RT "Identification of a new type of mammalian myosin heavy chain by
RT molecular cloning. Overlap of its mRNA with preprothymosin B
RL J. Biol. Chem. 262:14625-14632(1987).
RN [2]
RP SEQUENCE OF 1-789 FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=92268028; PubMed=1587791;
RA Kawakami H., Moriyoshi K., Utsami T., Nakajishi S.;
RT "Structural organization and expression of the gene for bovine myosin
RT I heavy chain."
RL J. Biochem. 111:302-309(1992).
CC -1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
CC ACTIN FILAMENTS (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J02819; AAA30658.1;
CC PIR: A29483; A29483.
CC HSSP: P08799; 1NMD.
CC InterPro: IPR000048; IQ_region.
CC InterPro: IPR001609; myosin_head.
CC Pfam: PF00063; myosin_head.1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC ProDom: PD000355; myosin_head.1.
CC SMART: SM00015; IQ_3.
CC SMART: SM00242; MSc_1.
CC PROSITE: PS00096; IQ_3.
KW Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
KW Multigene family.
FT DOMAIN 1 681 MYOSIN HEAD-LIKE.
FT DOMAIN 2 697 719 IQ 1.
FT DOMAIN 3 720 742 IQ 2.
FT DOMAIN 4 743 772 IQ 3.
FT NP_BIND 101 108 ATP (BY SIMILARITY).
FT DOMAIN 5 571 593 ACTIN-BINDING (POTENTIAL).
FT CONFLICT 708 708 MISSING (IN REF. 2).
SQ SEQUENCE 1043 AA; 118868 MW; 581BE91EA508B3C CRC64;
Query Match 39.08; Score 2106.5; DB 1; Length 1043;
Best Local Similarity 44.08; Pred. No. 4.5e-126;
Matches 455; Conservative 185; Mismatches 330; Indels 65; Gaps 21;
Qy 27 VGVDFVLENTFSEAFIENLRKRRRENILTYIGVYVSNVPRYDLQIYSQHMER 86
Db 8 VGVDFVLENTFSEAFIENLRKRRRENILTYIGVYVSNVPRYDLQIYSQHMER 86
Qy 87 GVSEYVPPHLFVAVDVYRALTERRDQAVMISGESGAGKTATKRLQFVETCPAPE 146
Db 67 DYTFELPKPIYALANAYOSLNRDRDCIILITGESGAGKTATKRLQFVETCPAPE 126
Qy 147 RGAAYVDRLLQSNPVLAEFGNAKTNNDNSRRGKYMIDQDFKGAIPGGHILSTLLEK 206
Db 127 QVNSVKEOLLQSNPVLAEFGNAKTNNDNSRRGKYMIDQDFKGAIPGGHILSTLLEK 186

```

QY 207 RVVHNGHBERNHVYVOLLBEGEETLRRLGLERNPOSTLYLYKGCACVSSINDKDMK 266
 187 RYVQLEGERNFHFIFOLLAGADAOLKALKLERDGTGAYL- NPTSTSHVMDADDANK 245
 QY 267 VMKALSVIDEFEDEVEDLSTIVASVHLGHNFADDESN---NQVTENQKLTJTL 322
 246 VLOSAMTVIGFSDEROLEVALYALVGLGNVELINEFQANGVPASGIDGKGVGEL 305
 QY 323 LGVEGTLREALTHRTIKAGBELLSPMLEQAAYARDAIAKAVYSRPTWLVKINSL 382
 306 VGLNSVLEALALCSRTMETAKEKVYVTLTVIAQOVARDAIAKAVYSRPTWLVKINSL 365
 QY 383 ASGADSPSRKSTTVVGLDIYGFVFNHNSFQFCINCNELQOLFELIKSQDEY 442
 366 KVGTEG---KRVKGVLDIYGFELIENNSFEQFINCNELQOVLEMTLKEQEBY 420
 QY 443 EAGIAMEPYOYNNKTIICDLVEKFKIISIDECELRPGATDTLTLEKEDYKPP 502
 421 KRGGITWVAVYEDNGLICNLHNGRGLIAMDDECLRGVYSDSTFLAKNOLFESK 480
 QY 503 HPLTHKILADOKTRK---SLDRGEFLLHYAGEVYVSGFLDKNDLLEFNLKETMCSM 559
 481 HY-ESKVTQNAQROHSMGLSCFRIHYAGKVTYVNSFIDKNDLLEFRLSQAMMKAR 539
 QY 560 NPIMACOFDSE--LSDKRREPVTATOPFKMSILOLVEILRSKPAYIRICIKPDADKOP 617
 540 HPLRLSLFEPGDKQASLKRPPAGQOFSSVTLTKNLYSKPNATIRICIKPREHQORH 599
 QY 618 FDEVILRHQVYLGIMENLVRRAAGFRRKRYEALFQRYKSLCPEPMWAGRPDGVAV 677
 600 FSEFLVSVQAOYGLLENVVRAGAYRQAYSGFLERYRLLSRSTWPMNGGDQGVK 659
 QY 678 LVNHLCKPBEYKMGRTKIFIRPKLTFATEDSLVFRQSLAKTQANRGFWKQKFLR 737
 660 VIELSMSSSEELAFGKTKLFRSPKTLFLEBORRLLOOLATLICKYRGWCRHYOL 719
 QY 738 VKRSALICISMRGTGLRRKRAKRWAAQTIRRLIGFLIRHSR-----CPENAF 788
 720 MRSQOIVISSWFRGNQKRYKMKASALLIOFVGMKARKRYRFGSALLISNFI 779
 QY 789 FLDHVRASFLNLRQLP- NVLDTSNPP-----PALREASELRELCMKMKWKYCR 842
 780 YKSMVQ-KFLGLKRLNDLPPSPILDKWPSAPYKFTYANLELORLHO-----WK-CK 830
 QY 843 ----SISPEWKOOLQKAVASELFGKNDYPOSVRLFTSTRTGTEESPRVLOSIGSE 898
 831 KFRDQSLSPKQVEVLRKELCASELFGKKNASTYPOSVRLFPFGDIYIGQR-NPKLOKAGE 889
 QY 899 --PIQAVPVVYKDKGKPRPQOLLTPSAVVI--VEDAKVQRIIDYANLNGISVSSLS 954
 890 EGPLMAETVAVKYNRGNAKTSRLLLTGKGVITITDMKNQOAKTVIPLNSLAGVSTSEK 949
 QY 955 DSLEFLVAVQREDN-KQKGVVLOSADVIEFLTK---TALSADVN-----NINING 1001
 950 DGLFSLHLSEISSVSGSEFLVSEHVITELTRICRATDATOMLPVTVEEFSVKEK 1009
 QY 1002 GSIT--FAGGPGRGD 1014
 1010 GSLTVKVIQGGGG 1024
 Db
 QY 1002 GSIT--FAGGPGRGD 1014
 1010 GSLTVKVIQGGGG 1024
 Db
 RESULT 5
 MYIA_MOUSE
 ID MYIA_MOUSE STANDARD; PRT; 1107 AA.
 AC P46735: P70244:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin IA (Myosin I alpha) (MMI-alpha) (MMIA) (MTH-L).
 GN MYOIA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6; TISSUE=Cerebellum;
 RA MEDLINE=93194946; PubMed=844986;
 RA Sher E.H., Joyce M.P., Greene L.A.;
 RT "Mammalian myosin I alpha, I beta, and I gamma: new widely expressed
 genes of the myosin I family.";
 RL J. Cell Biol. 120:1405-1416(1993).
 [2]
 RP SEQUENCE OF 13-1107 FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA MEDLINE=93209370; PubMed=8458427;
 RA Kostlovsky J.S., Qian C., Jiang X., Mercer J.A.;
 RT "Molecular cloning of a mouse myosin I expressed in brain.";
 RL FEBS Lett. 320:121-124(1993).
 CC -1- FUNCTION: MOTOR PROTEIN THAT MAY PARTICIPATE IN PROCESS CRITICAL
 CC -1- TO NEURONAL DEVELOPMENT AND FUNCTION SUCH AS CELL MIGRATION,
 CC NEURITE OUTGROWTH AND VESICULAR TRANSPORT.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PROMINENT EXPRESSION IS SEEN IN THE BRAIN,
 CC LUNG AND LIVER. IT IS ALSO EXPRESSED IN THE HEART AND TESTIS. A
 CC HIGH LEVEL EXPRESSION IS SEEN IN VIRTUALLY ALL NEURONS (BUT NOT
 CC GLIA) IN THE POSTNATAL AND ADULT MOUSE BRAIN AND IN NEUROBLASTS OF
 CC THE CEREBELLAR EXTERNAL GRANULAR LAYER.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 5 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: L00923; AAA9800.1; -
 CC EMBL: X69987; CAA49604.1; -
 CC HSSP: P08799; 1MND.
 CC MGD: MGI:107732; Myoia.
 CC InterPro: IPR000048; IQ_region.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00603; myosin_head.
 CC Pfam: PF00612; IQ_5.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART: SM00015; IQ_5.
 CC SMART: SM00242; MISC; 1.
 CC PROSITE: PS0096; IQ_4.
 CC Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
 CC Multigene family; Alternative splicing.
 CC
 CC FT DOMAIN 704 727
 CC FT 728 749
 CC FT 750 778
 CC FT 780 807
 CC FT 808 837
 CC FT 837 858
 CC FT NP_BIND 108 115
 CC FT DOMAIN 592 599
 CC FT VARSPIC 791 819
 CC FT CONFLICT 260 260
 CC FT CONFLICT 304 304
 CC FT CONFLICT 307 317
 CC FT CONFLICT 333 333
 CC FT CONFLICT 546 546
 CC FT CONFLICT 571 571
 CC FT CONFLICT 742 742
 CC SEQUENCE 1107 AA; 128527 MW; FA244ESD2DE752A5 CRC64;
 Query Match 38.7%; Score 2090.5; DB 1; Length 1107;
 Best Local Similarity 41.7%; Pred. No. 5, 1e-125;
 Matches 463; Conservative 183; Mismatches 346; Indels 119; Gaps 19;


```

0y      1008 GGPGRDGIIDFTSGSELLTRKKNNLAVVA 1038
       |:::|::|::|::|::|::|::|::|::|::|
Db      1083 QGNQKNGSV-----TCKRKNRLLEVA 1105

RESULT 6
MYIA_RAT STANDARD: PRT: 1136 AA.
AC ID MYIA_RAT
AD 005096;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin IA (Myosin I alpha) (MMI-alpha) (Mmia) (Myosin heavy chain myr
   1).
GN MYOIA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley;
RX MEDLINE=93194945; PubMed=8449985;
RA Rupert C., Kroschewski R., Baehler M.;
RT Identification, characterization and cloning of myr 1, a mammalian
   myosin-I.
RL Cell Biol. 120:1393-1403(1993).
CC -1 FUNCTION: MOTOR PROTEIN THAT MAY PARTICIPATE IN PROCESS CRITICAL
   TO NEURONAL DEVELOPMENT AND FUNCTION SUCH AS CELL MIGRATION,
   NEURITE OUTGROWTH AND VESICULAR TRANSPORT (BY SIMILARITY).
CC -1 ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND C; ARE
   PRODUCED BY ALTERNATIVE SPLICING.
CC CC -1 SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1 SIMILARITY: CONTAINS 5 IQ DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
   or send an email to license@isb-slb.ch).
CC -----
DR EMBL; X68199; CAA48287.1; -.
DR HSSP; P08799; IMND.
DR InterPro; IPRO000048; IO_region.
DR InterPro; IPRO01609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 6.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 6.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 5.
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
   Multigene family; Alternative splicing.
FT DOMAIN 1 703 MYOSIN HEAD-LIKE.
FT DOMAIN 1 703
FT DOMAIN 704 729 IQ 1.
FT DOMAIN 730 750 IQ 2.
FT DOMAIN 750 778 IQ 3.
FT DOMAIN 780 807 IQ 4.
FT DOMAIN 808 837 IQ 5.
FT DOMAIN 837 866 IQ 6.
FT NP_BIND 108 115 ACTP (POTENTIAL).
FT DOMAIN 592 599 ACTIN-BINDING (POTENTIAL).
FT VARSPLOC 794 823 MISSING (IN ISOFORM 1B).
FT VARSPLOC 794 852 MISSING (IN ISOFORM 1C).
SQ SEQUENCE 1136 AA; 131917 MW; 69D6C2A84E9070E2 CRC64;
Query Match 38.5%; Score 2080; DB 1; Length 1136;
Best Local Similarity 40.6%; Pred. No. 2,4e-14;
Matches 463; Conservative 183; Mismatches 346; Indels 148; Gaps 19;
```

```

27 VGVDFVLENTSEAFIENIRRRREVLITTYIGPVLSVNPYRDQIYSROMEXR 86
15 IGVDFVLEPL-NEETFDNLKRRDNEIYIGSVYISVNPRLSLIYSPEKEDYR 73
87 GSEFEVPHLFAVADTVYRALRTERDOAVMISGESGAKTEATKLOFAETCPAPE 146
74 NNNFIELSHIRALDEAVRSILRDOCKOCLITGESGAKTEAKLWMSYAAACGGA 133
147 RGVAVDRLLQSNPVLEAFNAKTLRNDSSRFQYMDVDFKAPVGGHLLSYLKS 206
134 EYNVVEQLLOSTPVLEAFNAKTVRNDSSRFQYMDIEFDFKDPGLGVISNYLKS 193
207 RVYHONHGERNFHVYOLLEGEETIRLGLERHPOSLYLVKQCKVSSINXSKMK 266
194 RVYKOPRGRNHFVYOLLEGEETIRLGLERHPOSLYLVKQCKVSSINXSKMK 252
267 VNRKALVIDFEDVEDELLSIVASYLHGINIFADESDN---AQVTENOLKYLRL 322
253 TVRNMQIYGFSDPEASVLEVAVALKGNIEFPESMNGLDSKIKDKNELKEICEL 312
323 LGVEGTLREALTTHRTIKAGBELLSPLEBOAARALAAVSRFTWLVYKINSL 382
313 TSIDVYLERAFSFEVTEKOKVSTLVVAQYARALANKLSRLEFWLVNINSEI 372
383 ASKDAEPMRSTVYGLDIYGFVYOHNSFEOFCINCNKLOLFTELTKSEOEY 442
373 KQOTKVRK-----WGVLDIYGFELFENSFEOFTINCNKLOLFTELTKSEOEY 427
443 EAEGLAMEPVYFNKTIODLVEKFKGIISLDEBCLRPGEATDLTELEKEDTVKPPH 502
428 IREDIEMTHIDYFNNAICDLIENNTNGILAMDECLRPGEATDETELEKINQVATHQ 487
503 HELTH--KIADOKTRSLDGEPRILHAGEVYVSTGTLNNDNLINKLETCSSNN 560
488 HESRNSKCSRLINDTLPHSCRIODHAGVLYOEGVDKNDNLIDLSQAMKACH 547
561 PMAOCFDSKSELS--DKRPETVATOPFKKSLQVLEILSKSPAYIRCIKPDANKOPRF 618
548 ALIKSLFPBSNPAKVALKRPPTAGSOFVATIMKNLQTKNPNYRCIKPDKKAHIF 607
619 DEVLIRHOVYVIGLIMENLVRRAGVYRRKYEAFLORYSLCPETVPMAGRPDGVAVL 678
608 SESLYCHOIRYIGLLENVRRAGVAFROAYEPCLEKRYMLOKOTPMHKKGPARSVEVL 667
679 VRHLGKPEYKMGRTKIFIRPKTIFAEEDSLVROSLARQIAOKAFHROKELFV 738
668 FNELETPVEYSGSKIFIRPKTIFAEEDSLVROSLARQIAOKAFHROKELFV 727
739 KRSALCIQSWMGRTGLRRRAAKKWAQTIIRLIRG-----ILRHSRCPENA--- 787
728 KRSQVIAAMVRYAOQKRYOQIKSSALVYISYIRGMARKILRELKHOKCKEATTTIA 787
788 ----- 787
788 AYWHGTQARKERRRLKDEARNKHAIVIAFWLGSARRELKLEAKRNKHAIVIAV 847
788 -----FLDHYRASFLINLRLOLP- NVLDTSMPTPPALR 822
848 WLGKAVRREKFFRANACKKIYEFLOIIVKYLLENMKNKPSLIDKNNPSRYLTL 907
823 EASELLRELCKMKNV---KYCRSISPEMKOOLQKAVASEIFPKGKNDYPOSVPLFIS 879
908 DSH- - - - -KELKRIPIHWRCKKYRDOQKLIYEELKSESELEFKKALYPSVSGPFOG 965
880 TRGTETELSP-----RVLOSISGPLOYAVPVYKDYDKGKVPRRQLLPPSAVYIE--D 933
966 AYI---EINKNPKYKLLKDAIEKIIIAEVYNKINANGKSTSRITLTLNNNLADQKS 1022
934 AKYKORIDYANLIGISVSLSDLFYLVOR--EDNKOKDVLVLOSQSHVLETTLK--TAL 989
1023 GQIKSEYPLVDYKVMSSONOGFVHLKESSEASKODFLFSSDHLIEMATKLYRTL 1082
990 SADRVN-NINI-----NGSITFAGGPGRDIIDFTSGSELLITKAKNGLAVYA 1038

```

```

DB 1083 SOTKOKNIEISDEFLVQFRDQKVCYKFIQGNKNSVP-----TCKRKNRLLEVA 1134
RESULT 7
MYHL_HUMAN
AC Q9UBC5; Q9UBD7; STANDARD; PRT; 1043 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brush border myosin I (BBM-I) (Myosin I heavy chain) (MHC).
GN MYHL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RA Bixie D.D., Munson S.J.;
RL "Human brush border myosin I";
RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Jejunal brush border;
RX MEDLINE=20119970; PubMed=10653589;
RA Li W., Wang J., Coluccio L.M., Matsudaire P., Grand R.J.;
RL "Brush border myosin I (BBMI): a basally localized transcript in human
jéjunal enterocytes.";
RN J. Histochem. Cytochem. 48:89-94(2000).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99073867; PubMed=9858156;
RA Skovron J.F., Bement W.M., Mosseker M.S.;
RL "Human brush border myosin-I and myosin-Ic expression in human
intestine and Caco-2Bb cells.";
RC Cell Motil. Cytoskeleton 41:308-324(1998).
-1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
ACTIN FILAMENTS (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC EMBL: AF127026; AAD31189.1; -
CC EMBL: AF105424; AAC78645.1; -
CC EMBL: AF009961; AAC27437.1; -
CC HSP: P08799; IAMD.
DR Genew; HGNC:7595; MYO1A.
DR InterPro; IPR001048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 3.
DR PRINTS; PR00193; MYOSINHEAVY.
DR Prodom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 2.
KW Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
KW Multigene family.
FT DOMAIN 1 681 MYOSIN HEAD-LIKE.
FT DOMAIN 697 719 IQ 1.
FT DOMAIN 720 742 IQ 2.
FT DOMAIN 743 772 IQ 3.
FT NP_BIND 101 108 ATP (POTENTIAL).
FT FT ACTIN-BINDING (POTENTIAL).

```

FT CONFLICT 446 446 Q -> P (IN REF. 3).
 FT CONFLICT 712 712 R -> G (IN REF. 3).
 SQ SEQUENCE 1043 AA; 118400 MM; EESD79D9DE85D0C CRC64;

Query Match 38.28; Score 2064.5; DB 1; Length 1043;
 Best Local Similarity 44.5%; Pred. No. 2.1e-123;
 Matches 440; Conservative 186; Mismatches 320; Indels 43; Gaps 19;

27 VGVDFVLENTFSEAFETENLRREMLITYIGPVLVSNVPRDLOIYRQHERR 86
 8 VGVDFVLENTFSEAFETENLRREMLITYIGPVLVSNVPRDLOIYRQHERR 86
 87 GVSFEVPHLFAVDYVYRALTERRDQVMSISGSGAGKTEATKRLQFAETCPAPE 146
 67 DYFVELKPHIYALAVAVOSLRDRDQCILITGSSGSKTEASLVSYAAVCGKE 126
 147 RCGAVDRLLQSPVLEAFGNAKTLANDSSRFGKYMDVQDFKAPVGGHLLVLEKS 206
 127 QVNSVEQOLQSPVLEAFGNAKTLANDSSRFGKYMDVQDFKAPVGGHLLVLEKS 186
 207 RVVHONGERNPHVFEVQLEGGEEELRLGLEENQSLYLKGCACAVSSINDKSPK 266
 187 RLKQKLGKERNHIFQLAGADEQLKALKLEEDTGYALV-NHEVSHVDMDDASSER 245
 267 VMKALSVIDEFEDEEDLLSIASVYLGNIHFADEDSNAQVTE-----NOLKYLTR 321
 246 AVQSAVAVIGFSEEEIROYLEVTSWVLKIGNV-LVADEQASIGAPSGIRDRGREIGE 304
 322 LGVEETTLREALTHKTIKAGBELLSPNLDOAAVADALAKAVYSRTFTLVKRNKS 381
 305 MGLNSEEVEERLACSTMETAKEKVYALNVQAQYARDAKLNKNTSRLEFDMIVRNES 364
 382 LASKDAPSMRSTYVGLLDIYGFVFNQHSPEOCINVCNEKIQLEFTELTLSEDEE 441
 365 IAVGIGEKK-----VMGVLDIYGFLEEDNSFEQVINYCNKIQLEFTELTLSEDEE 419
 442 YEAEGLAMEPVQYFNKKIICLVEEKFGIISILDECLRPEDATDLEKLEEDTVKPH 501
 420 YKREGIPMTKVDYFDNGIICLLEHNOGILAMLEDECLRPVDSPTFLANKLDFSKH 479
 502 PHFLTHKLADQTKRSLDR-----GERLLHVAQETVYSTGDLKNNLDFLNKLETMS 557
 480 GHT--ESKVTQNAQROYDHTMGLSCFRICHYAGKYVNTSTSIDKNNLDFLQAMWK 537
 558 SMNPWACCFDKS--ELSDKKRPETVATQFKMSLLQVLEILSKPEAVIRCIKPDAPQ 615
 538 AQHPILRLSLFPPGNPKQASLKRPPAGAFKSSVALIMKNLISKSPNITKIKPNEHQOR 597
 616 GRPDEVYLIRHOVYVGLMELNLRRAAGFAVRRKYEAFLORYKSLCPETPMWAGRPDQV 675
 598 GOFSSDPLVATQARYGLLENVAVRRAGVYAHROGYGPFLEERYLLSRTPMNGDREGV 657
 676 AVLVRHLGYKPREYKMKRTKIFIRPKTIFATEDSLFVARQSLAKRQIQAAMGFIMROKE 735
 658 EKVVLGELSSGELAFGKTKIFIRSPKTLFYLEEORRLQLQALDILKIFYGMRCRHY 717
 736 LNVKSAICIQSMWMTGLRRRAAKRKMAOTIRRLIRGLRHSR--CEPNAFL-D 791
 718 QLMRSQILISSMFGNMOKKCYGKIKASVLLQAFVGRMKRKRRTKIFSEALITLAD 777
 792 HVRRAS---FLNLRRQDP-RNVLDTSWETP-PALREASSELLRELCKMNVKXCR-- 842
 778 FTKYSWVQKFLGLTKNNLPSTNVLTWPAFYKCLSTFANQLOLFQ--WKCKRPD 834
 843 SISPEMKOOLQOKAVASEIFEGKKKNYPSVPRLLISTRLGHEEISPRVLSLGE--PI 900
 835 QLSPOVELTRKELCASLEFGKKKASVPSVPIPCGDYIGLG--NPKLQKLKGGEEGPV 893
 901 QYAVPVYKRGYKPRROLDTLSAVVYEDAVKOR--IDVNLIGISVSDSLF 958
 894 LMAEAVKKNVRCNGTSSRILLTLTGHYILDTTKKSQAKIYIGLDNVAGSVTSLKDLGF 953
 955 VLVHOREDN-KQKGVVLIQSDHVIETLTK 986

Db 954 SLHLEMSVSGSGDFLVSEHVIETLTK 982

RESULT 8
 MYHL-CHICK STANDARD; PRT: 1045 AA.
 AC P47807: 090573;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Brush border myosin I (BBM-I) (BBMI) (Myosin I heavy chain) (MHC).
 GN MYHL.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66114119; PubMed=8681398;
 RA Collins K., Matsudaira P.T.;
 RT "Recombinant expression of the brush border myosin I heavy chain.";
 RL Cell Motil. Cytoskeleton 32:151-161(1995).
 [2]
 RP SEQUENCE OF 46-1045 FROM N.A. AND SEQUENCE OF 332-343 AND 490-503.
 RX TISSUE=Intestine;
 RX MEDLINE=90078325; PubMed=2687288;
 RA Garcia A., Coudrier E., Carboni J., Anderson J., Vandekerckhove J.,
 RA Moosker M., Louard D., Arpin M.;
 RT "Partial deduced sequence of the 110-kD-calmodulin complex of the
 RT avian intestinal microvillus shows that this mechanoenzyme is a
 RT member of the myosin I family.";
 RL J. Cell Biol. 109:2895-2903(1989).
 [3]
 RP 3D-STRUCTURE MODELING OF 699-731.
 RX MEDLINE=97148341; PubMed=8994973;
 RA Houdusse A., Silver M., Cohen C.;
 RT "A model of Ca(2+)-free calmodulin binding to unconventional myosins
 RT reveals how calmodulin acts as a regulatory switch.";
 RL Structure 4:1475-1490(1996).
 CC -1- FUNCTION: COULD PLAY AN IMPORTANT ROLE IN MORPHOGENESIS AND
 CC FUNCTION OF INTESTINAL MICROVILLI.
 CC -1- TISSUE SPECIFICITY: INTESTINE.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 CC EMBL: U04049; AAB38373.1;
 CC EMBL: X58479; CA41388.1;
 CC PIR: A33620; A33620.
 CC PDB: 1AUI; 17-SEP-97.
 CC InterPro: IPR000048; IQ_region.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00063; myosin_head.1.
 CC Pfam: PF00612; IQ.3.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC ProDom: PD000355; myosin_head.1.
 CC SMART: SM00015; IQ.5.
 CC SMART: SM00242; MYSC.1.
 CC PROSITE: PS50096; IQ.3.
 KW Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
 KW Multigene family; 3D-structure.
 FT DOMAIN 1 684 MYOSIN HEAD-LIKE.
 FT DOMAIN 701 727 IQ 1.
 FT DOMAIN 723 750 IQ 2.

Matches 406; Conservative 177; Mismatches 292; Indels 32; Gaps 12;

QY 27 VGVDFVLENTSEAFIENLRRENTLITYIGVYVSNPYRDLQIYSQHMERYR 86
 DB 8 VGVDFVLENTSEAFIENLRRENTLITYIGVYVSNPYRDLQIYSQHMERYR 66
 QY 87 GVSFEVPHLEFVADTVYRALTERRDQAVMISGSGAGKTATKLOFVETCPAPE 146
 DB 67 DYFEYELKPHIYALANAVQSLKDRDQOCILITGSGAGKTATKLOFVETCPAPE 126
 QY 147 RGAVRDLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 206
 DB 127 QVNSVKEQLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 186
 QY 207 RVVHONGERNPHVYFOLLEGGEEELRLRLGLERNOSYLYLYKGCARVSSINDKSDMK 266
 DB 187 RVVHONGERNPHVYFOLLEGGEEELRLRLGLERNOSYLYLYKGCARVSSINDKSDMK 245
 QY 267 VMKALSVIDFTEDEVEDLSIVASVYLHGNHFAADESDN---AQVTENQKYLTRL 322
 DB 246 AVQHANSVIGFSEETIRNOVLEATVYKLGKGNVLTDEFGANGIPASGIDGKIOIIGEM 305
 QY 323 LGVEGTLRALTHRIKIKANGELLSPNLQOAYADALAKAVYRTTWTYKRNRSI 382
 DB 306 MGLNSTELERALSRTMETGKEVYVYLVANTQAYARDALAKAVYRTTWTYKRNRSI 365
 QY 383 ASKDAESPWRSTTVGLDIDYGFYFQHNSEFQFNCINCKEQLQOFLTELTKSQEEX 442
 DB 366 KVGTGEKK-----VGVLDIYGFELLEDNSFEQFVLANCNEKQOFLTELTKSQEEX 420
 QY 443 EAGCIAMEPVQYNNKICDIVERKGIISIDEBCLRGEATDITFEKLEDTYKPPH 502
 DB 421 KRGIPMTKYEYEDNIGLNLHSGRGLAMDECELRPGVYSDSTFLAKLQLEFSKS 480
 QY 503 HFTLHKLADOKTKRSDR---GEFRLHYAGGVTSYVGFQDNKNDLLEFRLNKEFMCS 558
 DB 481 HY---ESKVSOMARQYDRTMGSLCFRISHAGVYVNTVGFIDKNDLLEFRLNKEFMCS 538
 QY 559 MNPIMACDSELSDK--KRPETVATQFKMSLQLEVLIRSEKPAVYRICIRPNDAKOPG 616
 DB 539 OHPLKSLPEEGNPKFASLKRPTACTOFKNSAVAMKMLYNSNPYICIRPNDAKOPG 598
 QY 617 RPEVILIRHOVKTLGIMENIRVRAAFYRKYKFAFLQYKSLCPETPMAGRPDQVYA 676
 DB 599 RPEVILIRHOVKTLGIMENIRVRAAFYRKYKFAFLQYKSLCPETPMAGRPDQVYA 658
 QY 677 VLVRLHGYKPEEYKMGRTKIFIRFPTLATESLEVRROSLATKIQAMRPFHMKOKL 736
 DB 659 KVLGSITLSEELAYKTKIFIRFPTLATESLEVRROSLATKIQAMRPFHMKOKL 718
 QY 737 RVKRSALICIQSWMRGTGLGRKAKRMAQITRLIRGLIRHSPRCPENA-----F 788
 DB 719 OMKRSQILISAMFRGKQKHYKIRSSVYLLQAFRGRAKRNRYKRSAGALTLANF 778
 QY 789 FLDHVASFLNLRRLP--RNYLDTSWPTPP--PALREASELLRELCMKMMWKYCRSISP 846
 DB 779 IYOSMAQKFLNKKLPSTKVLNDNTWPAAPYRCFNTANQLORLFYOKKCKKFRLOSLP 838
 QY 847 EKKOOLQKAVASELFKGRKNDYPOVPRILFSTRLGEISPRVLOSIGSE--PIQVAY 904
 DB 839 KOVOTLREKLASSELFGKKKASYPOVPRILFSTRLGEISPRVLOSIGSE--PIQVAY 897
 QY 905 PAVKIDR 911
 DB 898 TVKKNR 904

RESULT 10
 MYHL_RAT
 ID MYHL_RAT STANDARD: PRT: 842 AA.
 AC 062774;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Brush border myosin I (BBM-1) (BBM1) (Myosin I heavy chain) (MHC)
 GN MYHL OR BBM1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI:TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver, and Testis;
 RX MEDLINE=95298044; PubMed=7779104;
 RA Balish M.F., Coluccio L.M.,
 RT Identification of brush border myosin-I in liver and testis. J.
 RL Biochem. Biophys. Res. Commun. 211:331-339(1995).
 CC -1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
 CC ACTIN FILAMENTS (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U25148; AAA89132.1; -.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ_3.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ_3.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ_2.
 DR Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat;
 KM Multigene family.
 FT NON_TER 1
 FT DOMAIN 1 673 MYOSIN HEAD-LIKE.
 FT DOMAIN 689 712 IQ 1.
 FT DOMAIN 713 733 IQ 2.
 FT DOMAIN 735 764 IQ 3.
 FT NON_TER 842
 SQ SEQUENCE 842 AA; 97210 MW; DDAFDCB663083168 CRC64;

Query Match 35.8%; Score 1932.5; DB 1; Length 842;
 Best Local Similarity 45.6%; Pred. No. 3.7e-115;
 Matches 388; Conservative 165; Mismatches 269; Indels 29; Gaps 10;

QY 28 VGVDFVLENTSEAFIENLRRENTLITYIGVYVSNPYRDLQIYSQHMERYR 87
 DB 1 GVEDVILLEPL--DEESLKNLQURKKEKTEIYIGVYVSNPYRDLQIYSQHMERYR 59
 QY 88 VSEFEVPHLEFVADTVYRALTERRDQAVMISGSGAGKTATKLOFVETCPAPE 147
 DB 60 YTFELKPHIYALANAVQSLKDRDQOCILITGSGAGKTATKLOFVETCPAPE 119
 QY 148 RGAVRDLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 207
 DB 120 VNSVKEQLQSNPVLEAFGNNAKTLRNDSSRFKGYMDQDFKAPVGGHLSYLEKS 179
 QY 207 RVVHONGERNPHVYFOLLEGGEEELRLRLGLERNOSYLYLYKGCARVSSINDKSDMK 267
 DB 180 RVVHONGERNPHVYFOLLEGGEEELRLRLGLERNOSYLYLYKGCARVSSINDKSDMK 238
 QY 267 VMKALSVIDFTEDEVEDLSIVASVYLHGNHFAADESDN---AQVTENQKYLTRL 323
 DB 246 AVQHANSVIGFSEETIRNOVLEATVYKLGKGNVLTDEFGANGIPASGIDGKIOIIGEM 298

QY 324 GVEGTTLREALTRKRIIANGCELLSPLNLEQAAAYARDALAKAYSRFTWLVKIRNSIA 383
 DB 299 GINSVELERALSRTETGKEKAVYLVNIOAYADALAKNIYSLEFMYIRKIESIK 358
 QY 384 SKDASPSWSTVILLDIYGEVFOHNSFEQFCINCEKIQLOQFIETLKSEQEEYE 443
 DB 359 VGGE-----KRYMGVLDIYGEFIELDENSFEQFVINCENKIQOYFIELTKEQEEYE 413
 QY 444 AEGIANEPVOYFNNKIIICDLVEKFGIISILDECLRGGEADLTFLEKLEETVPHH 503
 DB 414 REGIPYKVEYFNGIICDLEHSORGILLAMDECLRGVSDTFLEKLQNFESKSH 473
 QY 504 FLTRKADQKTRSLR---GEFRLHAGEVYVGTGLDKNNDLFRNKETKSCSSM 559
 DB 474 Y--ESKYTONAROYDRIMLSOFRICHTAGVYTVTSIDKNNDLFRDLSQTMWMAO 531
 QY 560 NPIMAQCFDKS--ELSDKRPFEVAFQFMSLIQVLEILRSKREPAVIRICIKPDANOPGR 617
 DB 532 HPLHSLFPRGNPKESAPKRPAGTOFKNSVALINKNLYSKNPYIRICIKPDQOOGGR 591
 QY 618 FDEVLRHOYKVIYGLIMENLVRRAFGAYRRKYEAFIORYKSLCPETMPMAGRPDQVAV 677
 DB 592 FTSELVWQARYIGLEENVRAGAFRAVYFLEIRKILSRSTWPRNGDREGVEK 651
 QY 678 LVHLAGYKPEEYKMGRTKIFIRPKFLATEDSLLEVRSOLATKIQAMRGFHWKQFLR 737
 DB 652 VLCSLISSELAAYGRKIFIRPKFLATEDSLLEVRSOLATKIQAMRGFHWKQFLR 711
 QY 738 VKNSAICTQSMWGTGTCRKAARKNAQOTIRLNGF-----ILRSPCPENAF 789
 DB 712 MKRSQILSAMFGRGNKOKHKGKIRSSVLLIOAFVGRWKRKNRYRFGSGATITANLF 771
 QY 790 LDHVRASFLNLRROLF--RNVLDTSWPTP--PALREASSELRECKNMWKCRCRSISPE 847
 DB 772 YOSTOKFLINKNIPSTVLDNTWPAFYPGCPNTANQELQHLFQWCKKTRIDOLSPR 831
 QY 848 WKQOLOQKAVA 858
 DB 832 QVOTLRKICA 842
 RESULT 11
 MYIA.DROME STANDARD; PRT; 1011 AA.
 AC 023978;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE MYOSIN IA (MIA) (Brush border myosin IA) (BBMIA).
 GN MYO31DF OR CG7438.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RX MEDLINE=94260541; PubMed=8201616;
 RA Morgan N.S., Skovronsky D.M., Attavanis-Tsakonas S., Mooseker M.S.;
 RT "The molecular cloning and characterization of Drosophila melanogaster
 myosin-IA and myosin-IB."
 RL J. Mol. Biol. 239:347-356(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Abgayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brötter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davernot L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [3]
 RP FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
 STAGE.
 RX MEDLINE=96069904; PubMed=7598814;
 RA Morgan N.S., Heinzelman M.B., Mooseker M.S.;
 RT "Characterization of myosin-IA and myosin-IB, two unconventional
 myosins associated with the Drosophila brush border cytoskeleton."
 RL Dev. Biol. 172:51-71(1995).
 CC -1- FUNCTION: INVOLVED IN DIRECTING THE MOVEMENT OF ORGANELLES ALONG
 ACTIN FILAMENTS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. PROTEIN SHIFTS FROM THE
 BASOLATERAL TO APICAL DOMAIN IN ENTEROCYTES AND FOLLICLE
 CELLS.
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED IN GASTRIC CAECAE,
 MIDGUT CELLS OF THE PROVENTRICULUS, AND IN THE MID AND HINDGUT. IN
 THE LARVAL GUT BRUSH BORDER, EXPRESSION IS IN THE TERMINAL WEB
 DOMAIN. IN THE ADULT GUT BRUSH BORDER, EXPRESSION REMAINS IN THE
 WEB DOMAIN AND HAS ALSO MOVED INTO THE MICROVILLI. ALSO EXPRESSED
 AT LOW LEVELS IN FOLLICLE CELLS DURING OOGENESIS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
 THROUGHOUT DEVELOPMENT TO ADULTHOOD WITH HIGHEST LEVELS AT THE END
 OF LARVAL DEVELOPMENT. EXPRESSION IN EMBRYOGENESIS IS CORRELATED
 WITH THE FORMATION OF A BRUSH BORDER WITHIN THE ALIMENTARY
 CANAL.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 TO DOMAINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 DR EMBL: U07595; AAA19590.1;
 DR EMBL: AE003628; AAF52966.1;
 DR HSSP: P08799; IMND.
 DR FlyBase: FBgn0011673; MYO31DF.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR001609; myosin_head.

DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ; 2.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD00035; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS00096; IQ; 1.
 DR Myosin: ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
 KW Multigene family.
 FT DOMAIN 1 694 MYOSIN HEAD-LIKE.
 FT DOMAIN 695 721 IQ 1.
 FT DOMAIN 721 736 IQ 2.
 FT NP_BIND 100 107 ATP (BY SIMILARITY).
 FT DOMAIN 567 589 ACTIN-BINDING (BY SIMILARITY).
 SQ SEQUENCE 1011 AA; 117094 MW; 31206C065B5D50FA CRC64;

Query Match 33.4%; Score 1800.5; DB 1; Length 1011;
 Best Local Similarity 39.6%; Pred. No. 1.2e-106;
 Matches 426; Conservative 169; Mismatches 341; Indels 139; Gaps 29;

28 GVQDFVLENTSEAFLENRRLRRRLNLTITGCVLVSNPYRDLQISROMERYG 87
 8 GVQDFVLDVQVSMKFMNDLNRKFNQNSITTYIGVCVSNMPPRQNMITYPETIRKYG 66
 88 VSFYEPPLFAVDVYRALRTERDQAVMISGSGAGTETARKLQFYAETCPAP-- 145
 67 RELFENAPHLFAIDASAVRLKQRODCLISGSGAGTETARKLQFYAETCPAP-- 126
 146 ---ERGCANRDLQSNPVLEAFNAKTALNDNSREGKYMDVOFDFKGAIVGHHISYL 202
 127 NEIER--VKVYLQSNALILETFGNAKTRNDNSREGKYMDIEFDYKADPVGGITTYL 183
 203 LEKSNVHONGERNRHFVYOLLGEGEETLRGLGLENPQSYLYVQCQAKYSSINDK 262
 184 LEKSNVYQOQGERFHFSTQQLKAGANDNEKQYELQETKRYHLLNG--SMDILTEK 240
 263 SDMKVRRKALVIDTEDEVEDLSIVASVHLGNIHNAEDSDNAQYTTENQKYLTRL 322
 241 SDYKGTCAFNKFLGSTDVQTIWNTIAVHLGVEQTIED--ELVISNQHKLSTAKL 299
 323 LGVETTLREALTHKRIIAKEBELISPLNQAAYARALAKAVYSKRTFWLVKINSL 382
 300 LQVETELSTALTRKIVAGNVQKDMATQAEKQALAKAIYDLFTIIRIRINAI 359
 383 ASKDAESPMSSTVYLLGDIYGEVFNQNSFEQFCINCEKIQOQFTELTSEOEY 442
 360 LFRGSKTQA-RFNSTVIGLDIYGEFIPDSNFEQFCINCEKIQOQFTELTSEOEY 418
 443 EAEGIAWEPVOYFNKKIICDLVEKEKGIISILDECLRPGEATDLFLEKLEDTYKPH 502
 419 QREGIEMTNIEYFNKKIICDLVEQPHKGIIMDEACLSVGKVTDDTLGAMDKNLSKHP 478
 503 HPLTKLADOKTRSL-DRGEFRLHVAQGYTVGFLDKNNDLFR-----NLKET 554
 479 HTTSQQL--KPTDELKREDFRTHAGDVIVINGTEKNNKTLQODEFRLLHNSDA 536
 555 MCSNNPIMAOCFKSELSDKRPETVATQFKMSLQVLETLRSEKPAVYICIKPNDAQ 614
 537 NLSEKMPRGADIKT-----TKRPLTACTLFQRMADLVYLLKKEPPYVACIKPNDAK 592
 615 PGRDEVLIRHOKYGLGMLNLRVRRAGFAVRRYEAFLQRYKSLCPETWPM-AGRPD 673
 593 STVEDEEVEHQVRLGLENLRVRRAGFAVRRYEAFLQRYKSLCPETWPM-AGRPD 652
 674 GVAIVVRLHGLKPEYKMGRTKTRPPTLFAEDSLFVRRQSLAKIQAMGFIRM 733
 653 GVRVLEEKTF-ADQVKYGHKTRTSRTLFALEHONENIPIHIVLLQVRVGVIVRR 711
 734 KFLVRSALICIQSWNRGTLGRRAAKRMAQOTIRLLIRFIIIRSPRCENAFILDHV 793
 712 NFKKK-----AAITIVRAVYKVKLR----- 732
 794 RASFLLNLRQL--PRNVLD-----TSWTPPPALREASSELLR-----ELCKMMV-WKYCRS 843

DB 733 --STVOELANLRKAKMRDYGKSTQMPQPLAGRKVEAKLHRFEDWRAMMLTHKYPRS 790
 QY 844 ISPEWKOOLQOKAVASEIFKQKDNYPQSVPRLEIFSRLEETISPRVQSLGS----- 897
 DB 791 ---EW-PQLRLQIATAALAGRRRYWGA--RRWGYGLANSQENSGEYVNSIKNIN 844
 QY 898 -----EPQYAVPVYKDRGKYPKRPOLLTPSAVAVVEDAK-----VKORIDYAN 944
 DB 845 HPADGETFOOVLESFVKKNHFN-KQANRAFIYSDSTIHKLDGINKKFKDMKRTIKIRE 903
 QY 945 LTGISVSLDSLFVLHVQREDNKQGDVY--LQSDVITELTKTLASDRVN----- 995
 DB 904 LTISVSPGRQQLVFA---SSRKN-DLVFSLESE-----TLKEDRIEYGVGYVC 951
 QY 996 -----NININGSTIFAGCPGRDGIIDFTSGSELLITKANGHLAVAP 1039
 DB 952 KKYHDLGTLELVAVNTINISCRDQKARITVEASVVEPNRPREGNIIEFV 1006

RESULT 12
 ID MYSE_DICDI STANDARD; PRT; 1003 AA.
 AC 003479;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Myosin IE heavy chain.
 GN MYOE OR DMIE.
 OS Dictyostellum discoideum (Slime mold).
 CC Eukaryota; Mycetozoa; Dictyostellidae; Dictyostellum.
 OX NCBI_Taxid=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RX MEDLINE=93277957; PubMed=8504170;
 RA Urrutia R., Jung G., Hammer J.A. III;
 RT "The Dictyostellum myosin IE heavy chain gene encodes a truncated isoform that lacks sequences corresponding to the actin binding site in the tail."
 RL Blochim. Biophys. Acta 1173:225-229(1993).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE ACTIVITY THAT IS ACTIVATED BY ACTIN. MAY PLAY A ROLE IN MOVING MEMBRANES RELATIVE TO ACTIN.
 CC -1- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: L06805; AAA33201.1; -
 DR PTR: S33760; S33760.
 DR HSP: P08799; LMND.
 DR DictyDb; DD01049; MYOE.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; myosin_head.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF00612; IQ; 2.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD00035; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS00096; IQ; 2.
 KW Myosin; Actin-binding; ATP-binding; Calmodulin-binding; Repeat; Multigene family.
 FT DOMAIN 1 693 MYOSIN HEAD-LIKE.

FT DOMAIN 694 722 IO 1.
 FT DOMAIN 716 745 IO 2.
 FT DOMAIN 796 1003 NON ALPHA-HELICAL, C-TERMINAL DOMAIN.
 FT NP BIND 101 108 ATP (BY SIMILARITY).
 FT DOMAIN 556 630 ACTIN-BINDING.
 FT DOMAIN 858 958 MEMBRANE-BINDING (POTENTIAL).
 SQ SEQUENCE 1003 AA, 114852 MW, BE758BEC035766F CMC64;
 Query Match 32.28; Score 1739; DB 1; Length 1003;
 Best Local Similarity 38.9%; Pred. No. 9.5e-103;
 Matches 416; Conservative 171; Mismatches 328; Indels 154; Gaps 29;
 28 GVDFVLENTSEAPFENLRRFRRENILTYIGVIVSVNRYRLOIYSRQMERRG 87
 9 GVDFVLENTSEAPFENLRRFRRENILTYIGVIVSVNRYRLOIYSRQMERRG 67
 88 VSEFVPPHLEAVADTVYRALRTERRDQAMVIGESGAGCTEATKRLQFY---AETCP 143
 68 RYKEMPPIYALANAYRSMROSQENOCYIISGEGAGCTEASKIMQFLFVSSQSP 127
 144 APERGAVDRLLQSNPVLEAFGNATLNRDNRSGKYMVDQFDKGAIVGCHILSYLL 203
 128 NGRISKMLIR-EQSNPLLEAFGNATLNRDNRSGKYMVDQFDKGAIVGCHILSYLL 186
 204 EKSRYVHONGENRHFVYQLEGGEEETRLRLRNPOSYLYKGGCAKYSINDS 263
 187 EKSRYVHONGENRHFVYQLEGGEEETRLRLRNPOSYLYKGGCAKYSINDS 246
 264 DKRVKRALSVIDFTEDEVEDLISVAVLHGINHPAEDSNAQVT---ENOLKY 318
 247 EFKIYVAMETLGLKESDONSIMRILAIHINIFEAEEAEBORTGTVYKVSDTSLAA 306
 319 LTFULGVEGTLEALHTRK---TAKGELL-SPLNDEAAVARALKAIVSRFTETL 374
 307 AASCLTKDQSLTALCYRISIGVGRKCSIVPMDCNAAASRLAALERLEFNL 366
 375 VRKINRSLAKDAEBSRSTVYGLDIYFEVFOHNSPEOCINVCNEKLOQLEFTEL 434
 367 VSKIN-----TLSTYHREGVYIGLIDYFEVFOHNSPEOCINVCNEKLOQLEFTEL 419
 435 LKSEDEYEAEGJAMEVOYFNKKIICDLVEEFKGIISLDECLAPGATLPLEKL 494
 420 LKSEDEYEAEGJAMEVOYFNKKIICDLVEEFKGIISLDECLAPGATLPLEKL 475
 495 EDYVKKPHFLTKLADQKTRKSLDGE---FRLLHYAGETVYVTEFLDKNNDLPRNLK 552
 476 CKPEKPHLOSYVS-----KDRSNDGTCFLKHAAGVTVYVGRFLDKNNDLPRNLK 530
 553 ETWCSSNPIMAOCFDEKSLSD-KKRPETVATQFKMSLQVLEILSKREPAVYRICKPND 611
 531 SSMQSSDPLVQGLFPTREPDSKRRPRTAGSQFRANMANLITTLACSPHYRICKSND 590
 612 AKQGRFDEVLIRHOKYKYLGLMNLVRRAGFAVRRKYLEFLORYSLCPEYPMAGRP 671
 591 NKAGVYDEDRDHRVRYGLLENNVRAGFAVRRKYLEFLORYSLCPEYPMAGRP 650
 672 QDGAVALVRLHGLGKPEYKMGKPIRFPKTLFATEDSLLEVROSATQIQAAMGFHW 731
 651 KQATELLQOHNDKEIRMGKTKVIRIPRTLFYEKEKELMPLRIVILQITWNGYRA 710
 732 KQFELVYKRSALCISQWMTGLGRRRAAKRKAQOTIRLLRGF-ILRHSPPCENAFEL 790
 711 RSKW-NORKAKIKIQLF-----RSRYRKM---FREIHRFAKVDARPOWQKQVF-- 757
 791 DHRASFLMLRROLPRNVLDTSMPPRPALRASSELRLCKMKNVWYCRISIP--EM 848
 758 -----WPKDPSILDRADVLTHTI---HNCWRAEKMTLSLGA 791
 849 KQOLOKAVASEIFKGGK-----DNYPOSVPLRITSTRLETEI 887
 792 QNHMRQKVMAYDLFHGKKKMDFRHRHDADYLEKPMWNOCKYVLAQNLF----- 841
 888 SPRVLOSGLSEPIQIYAVPVVKYDRKGYKPRROLITPSAVVIVEDAKVKOR--IDYAN 944

Db 842 -----STYGDPEVFLADYIVKPNKG-V-PQARGIVYTGNTNKKDKPNKYNKMGTPLYD 895
 QY 945 LTGISVSSLSDFLVH---OREDNKOG-----DVLQSDHYETLTKRLALS 990
 Db 896 VTSISPMADDFVLHCKAPQDFDLGLGNGEYANSEITTVYQQ---VLKTLGVKLS 952
 QY 991 ADVNNININ-----OGSIT-----FAGGPRGRIIDE 1018
 Db 953 VQFTSITVYNNARPGRSDTLTTPADINNDPKLISQFKKGKKEATIOF 1001
 RESULT 13
 MYTF_MOUSE
 ID MYTF_MOUSE STANDARD; PRT: 1099 AA.
 AC P70248;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin If.
 GN MYOLF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN-BALB/C; TISSUE-Cochlea;
 RC MEDLINE=97237053; PubMed=9119401;
 RA Crozet F., Amraoui A.E., Blanchard S., Lenoir M., Ripoll C., Vago P.,
 RA Hamel C., Fizesmes C., Levi-Acobas F., Depietis D., Mattei M.-G.,
 RA Weil D., Pujol R., Petit C.;
 RT "Cloning of the genes encoding two murine and human cochlear
 RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
 CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
 CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
 CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
 CC SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SPLEEN, EYE,
 CC BRAIN, LUNG, SMALL INTESTINE, TESTIS AND COCHLEA. BARELY
 CC DETECTABLE IN HEART.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X97650; CAA66251.1; -
 CC HSSP: P08799; 1MND.
 CC MGD: MG1:107711; Myolf.
 CC InterPro: IPR00048; IO.region.
 CC InterPro: IPR001452; SH3.
 CC InterPro: IPR001609; myosin_head.
 CC Pfam: PF00018; SH3; 1.
 CC Pfam: PF00063; myosin_head; 1.
 CC Pfam: PF00612; IO; 1.
 CC PRINTS: PR00193; MYOSINHEAVY.
 CC PRINTS: PR00452; SH3DOMAIN.
 CC ProDom: PD000066; SH3; 1.
 CC ProDom: PD000355; myosin_head; 1.
 CC SMART: SM00015; IO; 1.
 CC SMART: SM00242; MYSC; 1.
 CC SMART: SM00326; SH3; 1.
 CC PROSITE: PS0096; IO; 1.
 CC PROSITE: PS50002; SH3; 1.

KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; SH3 domain;
 Multi-gene family.
 FT DOMAIN 1 677 MYOSIN HEAD-LIKE.
 FT DOMAIN 697 722 IQ.
 FT DOMAIN 1042 1099 SH3.
 FT DOMAIN 579 589 ACTIN-BINDING (POTENTIAL).
 FT NP_BIND 110 117 ATP (POTENTIAL).
 SQ SEQUENCE 1099 AA; 125947 MW; 7D1592310C08F0 CRC64;
 Query Match 32.0%; Score 1726.5; DB 1; Length 1099;
 Best Local Similarity 38.0%; Pred. No. 6.7e-102;
 Matches 399; Conservative 159; Mismatches 297; Indels 195; Gaps 24;

QY 28 GVQDFVLENTSEAFETLRRLRRRLTYTGIVYVSNVPRDLOYSRQMERRG 87
 DB 18 GVDDVNVLLSQIT-EDAIYRNLRHKKRMDYIFTYIGSVLISNPKOMYFIDREIDLKYG 76
 QY 88 VSEYEPPLFVAVDVTYRALTERRDAVMSISGSGAGTETARKLLQFYAETCPAPER 147
 DB 77 AVQYEEPHIYALTDMYRNMLIDENQCVIISGSGAGTVAAYKIMYISKVSGGDK 136
 QY 148 GGAVDRLQSNPVLEAFGNKTLRNDSSRFKYMVDQFPGKAPVGHLSLLEKSR 207
 DB 137 VQHVADILQSNPLLEAFGNKTLVANNSSRFKYMVDQFPGKAPVGHLSLLEKSR 196
 QY 208 VVQNHGERNFHVEYQLLEGEEETLRRLGLERNPOSYLTVKGCACAVSSINDKSPKV 267
 DB 197 VVQNHGERNFHVEYQLLEGEEETLRRLGLERNPOSYLTVKGCACAVSSINDKSPKV 255
 QY 268 MKKALSTVDFEDEVEDLLSTVAVLHGINHPAEDSSNAQVTTENQKYLRLLEKSR 327
 DB 256 TLSAVQVGFQACQVLOLVAGILHGINISF-CEEGNVARVESVSLAPVALLIGDS 314
 QY 328 TLRRLALFHRIIAK---GELLSPULLEQAAVARDLAKAVSRPTTWVYKINSLSLA 383
 DB 315 GRLOKLTLSRKNDSSWGRSSSIDVTYLVNQAATRLDLAGLAFEDFLVEALINRAM- 373
 QY 364 SKDAESPMSRSTTVYGLLDYIGFEVQHNSEFOFCINCNKCNKLDQLELLEKSEOEYE 443
 DB 374 OKPOEYS-----IGVLDYIGFEVQHNSEFOFCINCNKCNKLDQLELLEKSEOEYE 427
 QY 444 AEGIMEPVQYFNKKITDLYVEKFK--GIISIDDECL---RPGEATDLEKLEDTV 498
 DB 428 QEGIMPTIEYFNKKITDLYVEKFK--GIISIDDECL---RPGEATDLEKLEDTV 487
 QY 499 KPHPLFLTKLADQKTRKSLDGRFRLHYAGEVYVSTGFLDNNDLLEFNLEKTCSS 558
 DB 488 GTHEHF-----NSWSAG-FVHHYAGKVSIDVSGFCERNRDVLESDLELMOSS 535
 QY 559 MNPIMAQCF-DKSELSDKRRREYATQFKMSLDLVELIRSKEPAYIRICIRPNDAKOPGR 617
 DB 536 DODFLRMLFPEKLNIDKGRPSTAGSKIKQANDLVSLKCTPHYIRICIRPNDAKOPGR 595
 QY 618 FDEVILIRHQQVYGLGMENTLVRKRGAFVARKYEAFLQYKSLCETPMWAGRPQDQYAV 677
 DB 596 WEESVVKHQVEYLGRENIRVRKRGAFVARKYEAFLQYKSLCETPMWAGRPQDQYAV 655
 QY 678 LVRLHGYKPEYKMGKRTKIFRPFKTLFATEDSLVRRQSLATYIQAAWGRGFMHRKFLR 737
 DB 656 LLRAVNMEDQYQKOSTVYFKNPESLFLLEMEKRDGAFRTIQAM-----704
 QY 738 VKRSATCIQSWWRGTLGRKAKARKMAQITIRLIRGLFIRHSPREDNAFLDHYVASF 797
 DB 705 -----RRHYAVRRY-----713
 QY 798 LLNLRLRQPLRVNLDTSWPTPPALREASSELLRELCMKNMVKKYCRSISPEKKOOLQKAV 857
 DB 714 -----EEMRE-----E 719
 QY 858 ASEIFGKKNYPOVAPRLFISTRLGTEISPRVLOSIGS-EPIQYAVPVYKDYKGYK 916
 DB 720 ASNLILNKKERRRNSINNFVGNVLGLEE-RPELRQFLAKERERDFADSVTKYRR-FKP 777

QY 917 RPRQLITPSAVYVEDAKY-----KORDYANLNGISVSSLSDSLFLYHNG 963
 DB 778 IKRDLITPCVAVITGTEKKRPGKGLREVLRKIDQALGVSLSLRQDFLL--O 835
 QY 964 REDNKKQKGVLDVLDHVIETLTFT---ALSADRVNNININOGSITF-----A 1007
 DB 836 EE-----AADFTLESIFTEPVSLLCKRFEENARRPLITFDILQFVKKKKG 885
 QY 1008 GGPGRGDIIDFTSGS-ELLITKAKNGHLAV 1036
 DB 886 GGGTRN--VYFSAGOVNLAVALKAGGKLTIT 913

RESULT 14
 MSB.DICDI
 ID MSB.DICDI STANDARD; PRT; 1111 AA.
 AC P34092;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin IB heavy chain.
 GN MYOIA OR DMIB.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_Taxid=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RA MEDLINE=89345628; PubMed-2762320;
 RA Jung G., Saxe C.L. III, Kimmel J.A. R., Hammer J.A. III;
 RT "Dictyostelium discoideum contains a gene encoding a myosin I heavy
 chain."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6186-6190(1989).
 RP [2]
 RP SEQUENCE OF 481-490; 656-666 AND 783-798.
 RC STRAIN=AX3;
 RA MEDLINE=93315475; PubMed-8325874;
 RA Jung G., Fukui Y., Martin B., Hammer J.A. III;
 RT "Sequence, expression pattern, intracellular localization, and
 targeted disruption of the Dictyostelium myosin ID heavy chain
 isoform."
 RL J. Biol. Chem. 268:14981-14990(1993).
 CC -I- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY ACTIN. MYOSIN IB MAY HAVE A ROLE IN
 CC CHEMOTAXIS AND AGGREGATION: IT COULD SERVE TO STABILIZE AND EVEN
 CC RETRACT CORTICAL STRUCTURES, SUCH AS PSEDOPODS AND LAMELLOPODS.
 CC -I- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
 CC AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.
 CC -I- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION JUST BENEATH THE
 CC PLASMA MEMBRANE IN THE ANTERIOR PSEUDOPOD AT THE LEADING EDGE OF
 CC THE CELL.
 CC -I- DOMAIN: TH 1 BINDS DIRECTLY TO ANIONIC PHOSPHOLIPID MEMBRANES;
 CC MYOSIN I COULD THEREFORE MOVE ACTIN RELATIVE TO MEMBRANES AND
 CC VICE VERSA. TH 2 AND SH3 BIND TIGHTLY TO F-ACTIN; THIS TOGETHER
 CC WITH THE NUCLEOTIDE-SENSITIVE SITE IN THE HEAD, ALLOWS SINGLE
 CC MOLECULES OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.
 CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M26037; AAA33229.1;
 CC PIR; A33284; A33284.
 CC HSP; P08799; 1MND.
 CC DictyDB; DD01047; myoA.
 CC InterPro; IPR000048; myoA_region.
 CC InterPro; IPR001452; SH3.

```

DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00018; SH3_1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00612; IQ; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRINTS: PR00452; SH3DOMAIN.
DR ProDom: PD000066; SH3_1.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00242; MYSC; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR Myosin: Actin-binding; SH3 domain; Multigene family;
DR Ghemotaxs: Phosphorylation.
FT DOMAIN 1 694 MYOSIN HEAD-LIKE.
FT DOMAIN 922 1052 TAIL HOMOLOG REGION 1 (TH.1).
FT DOMAIN 1053 1111 SH3.
FT NP_BIND 102 109 ATP (POTENTIAL).
FT DOMAIN 547 627 ACTIN-BINDING.
FT DOMAIN 951 1015 ASN-RICH.
FT MOD_RES 332 332 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 1111 AA; 124313 MW; CD6386F08DC3642P CRC64;

Query Match 31.4%; Score 1692.5; DB 1; Length 1111;
Best Local Similarity 36.4%; Pred. No. 9.9e-100;
Matches 384; Conservative 172; Mismatches 312; Indels 187; Gaps 21;

QY 28 GVDGVLENTSEAEFIEMLRRFRRENLIYTYIGVLYSVNRYRLOLIYRQHMERYRG 87
DB 10 GTDLVWLPK-VSEDEICENLKRKRYNDFTYTNIGVLLSVNPFNNNSGPDIFAYRG 68
QY 88 VSEVEVPHFAVDVYRALRTERRDQAVMISGSGAGTEKTRKLLQFYAETCAPER 147
DB 69 KHAQEVPHHYQLAESRYRKMKNQDENCYIISGSGAGTEAKLIMGVSAISSTER 128
QY 148 GGAVERDLQSNVLEAFNAKTLRNDNSRFEKTYNDVQDFEGAPVGGHILSYLLEKSR 207
DB 129 VEYKHYVLESNPLEFAFNKTLRNNSRFRGKYEIOFDKAGDVGKGIYNYLLEKSR 188
QY 208 VVHONHGRNFHYVYOLEGGEETLRGLERNPOSYLYLVGCAKYSINDKDMVY 267
DB 189 VVYONPGRNFHYVYOLLGRASQEKRDYLS-SPESTYLLNOSCYTDGINDVSDYAE 247
QY 268 MKRALSYDTEDEVEDLSTVAVLHGNHRADEDSNAQYTTENOLKYLTRLGVES 327
DB 248 VROAMDTIGLTAQDQSDIIRIVACVLHIGNIYFIEDDKYAAIYDPNALAELASMLCIDS 307
QY 328 TITREALTHKRIAKG-----EELSLPLNEQAAYADALAKAVYSTFTVLYRKIN 379
DB 308 ATLQNALLFVYVITVGAGAGNRRSTYVNPQNEQANGTRDALARTIYRMFESWLVEKYN 367
QY 380 RSLASKADSPSRSTVGLGLDIYGEVQHNPSFOFCINYNENKLOOLFTLTKSRQ 439
DB 368 QSLASY-YKSP---YQNVIGILDIRFGEIERKNGFEQFCINFNENKLOOFFILITKAED 422
QY 440 EEEVAGEIAPVQYFNKKIICDLVEKF-KGIISLDECL--RPGATDITLFEKLE 495
DB 423 EEEVREGIKWEPIKYEYFNQVLCILIECKSPGIFSLDDICSLHAGSTGCDKFLKKA 482
QY 496 DYKPRPHF--LTHKLADQKTRSLDGERFLHYGEVYIYVSTGLDKNNLLFNRLKE 553
DB 483 GIYDGLHWRGM-----GFAIHYAGEVYVEEYEGESDKRKLDFDLIE 528
QY 554 TMSQSNPIAAQCFDEKSELS-DKRPETVATQKMSLDQVELLRKREPAYIRICIPNA 612
DB 529 AIOCSKMPFLASIFNEBTGSLQKKRPTTAGFKITKISAGELMKLSCOTPHIKCIPNEI 588
QY 613 KQGRPEDEVLIRHVOYKILGIMENIRYRAGFAVRRKYEAFLOKYSICLCEETPMMA---G 669
DB 589 KKAADMENSVKHOQVYIGLLENVVRRAAGFAVRNFDKYLKRYKLLSTKTIWIMEWNG 648
QY 670 RPDQGAVALVRHLGYRPEEYKMGRTYIFIRFPTLFTATEDSELEVRQSLATKIQAAMRG 729
DB 670 RPDQGAVALVRHLGYRPEEYKMGRTYIFIRFPTLFTATEDSELEVRQSLATKIQAAMRG 729

```

```

DB 649 DAIEGCKTIFQDMNLEAGQWLGKTKVFRHPEVYFLEELADKKDFDCTAKIOKFR-- 706
QY 730 HWRQKFLRKRAICIQSMWRGTLGRRAAKRMAQIIRLRINGFILRHSPCEENAF 789
DB 707 MKRK----- 711
QY 790 LDHVRASFLLNRLQPRNVLDTSWTPPPALREASELLRELCKMNMVKKYCRSISPBMK 849
DB 712 -----K 712
QY 850 QOLQKAVASEIFKKNKNDYPOSVRLISIRLAGEE--ISPRYQSLSEPIQYAVPY 906
DB 713 HSELRQALAHMFKKERORSIDRKTSYDIFDENOFQGEAMQNAKKERERVFAATV 772
QY 907 VYDGRKGRPRROLLTPSAVVIYEDKKV-----ORDVANLGISVSSLSDSLF 958
DB 773 IKIDRRA-KQKYEAVLTDQALYFEKSIKKVYLHILRRGLREIGVISTLSDDVI 831
QY 959 VLVHVRD-----NKGQDVVLOSDHVIETLFTALS--ADRVNNININQSIYFAGPG 1011
DB 832 VHLDEHQVLENDKTEILLVYEFRAIGGSLNVQFSDRIN-----YTLKKGEO 883
QY 1012 RDGIIDFTSGSE---LITKAKNGHLAVAPRLNS 1043
DB 884 KE--ISFQSEQCPPLVVKGGKGLIGTASGLPS 916

```

RESULT 15

MSD.DICDI ID MSD.DICDI STANDARD; PRI; 1113 AA.

AC P34109; 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin ID heavy chain.

GN MYOD OR DMD.

OS Dictyostelium discoidium (slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

NCBI_Taxid=44689;

LN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 604-610; 733-742 AND 914-928.

RC STRAIN=AX3; MEDLINE=93315475; PubMed=8325874;

RA Jung G., Fukui Y., Martin B., Hammer J.A. III;

RT "Sequence, expression pattern, intracellular localization, and

targeted disruption of the Dictyostelium myosin ID heavy chain

isoform";

RL J. Biol. Chem. 268:14981-14990(1993).

CC -I- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE

CC ACTIVITY THAT IS ACTIVATED BY ACTIN. MYOSIN ID MAY HAVE A ROLE IN

CC CHEMOTAXIS AND AGGREGATION; IT COULD SERVE TO STABILIZE AND EVEN

CC RETRACT CYTOSOL STRUCTURES, SUCH AS PSEUDOPODS AND LAMELLOPODS.

CC -I- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY

CC AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS.

CC -I- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION JUST BENEATH THE

CC PLASMA MEMBRANE IN THE ANTERIOR PSEUDOPOD AT THE LEADING EDGE OF

CC THE CELL.

CC -I- DOMAIN: TH.1 BINDS DIRECTLY TO ANIONIC PHOSPHOLIPID MEMBRANES;

CC MYOSIN I COULD THEREFORE MOVE ACTIN RELATIVE TO MEMBRANES AND

CC VICE VERSA. TH.2 AND SH3 BIND TIGHTLY TO F-ACTIN. THIS TOGETHER

CC WITH THE NUCLEOTIDE-SENSITIVE SITE IN THE HEAD, ALLOWS SINGLE

CC MOLECULES OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.

CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

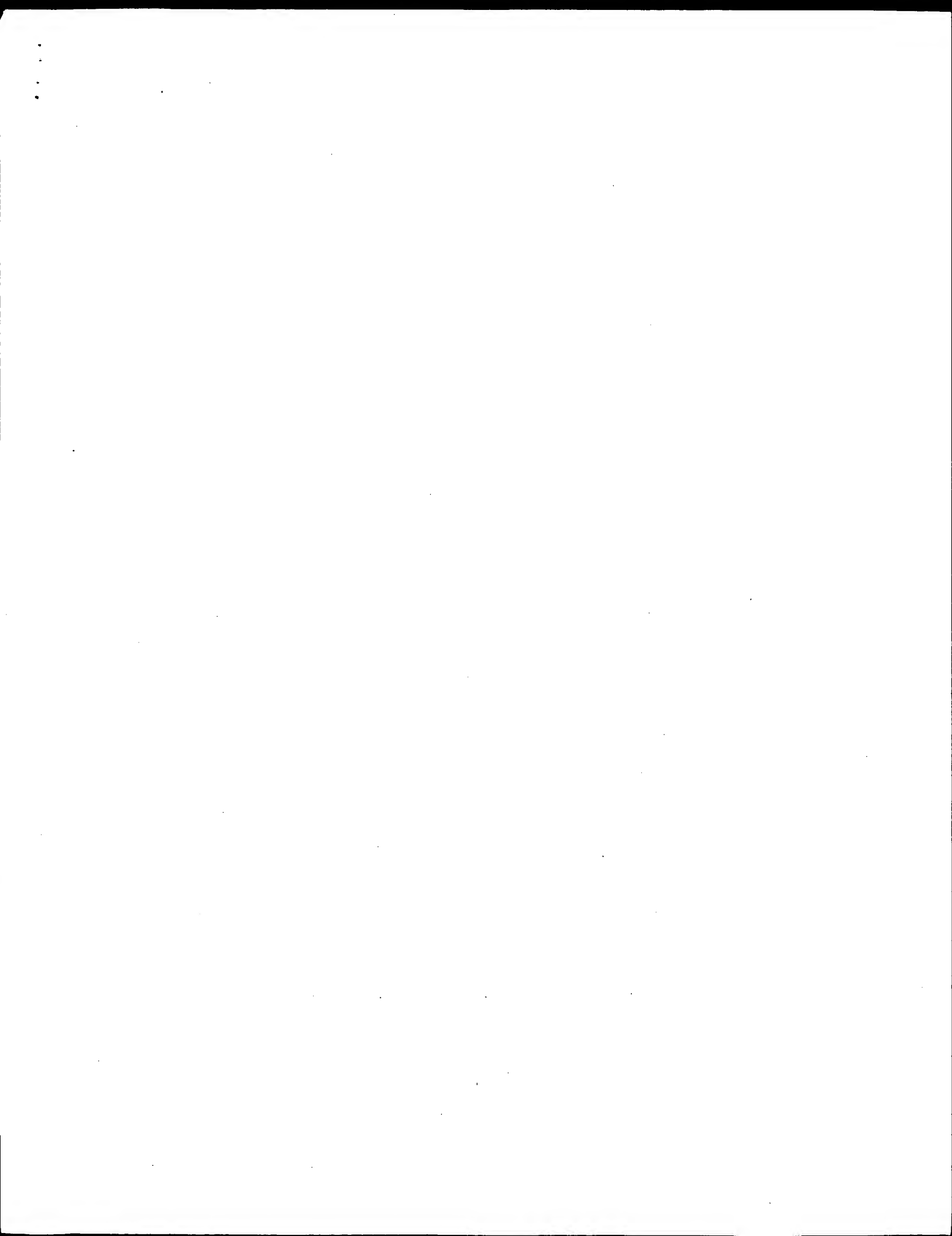
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR	EMBL: L16509; -; NOT_ANNOTATED_CDS.
DR	PIR: A47106; A47106.
DR	HSSP: P29354; IGRI.
DR	Dic1yD; DD01048; myoD.
DR	InterPro: IPR001452; SH3.
DR	InterPro: IPR001609; myosin_head.
DR	Pfam: PF00018; SH3; 1.
DR	Pfam: PF00063; myosin_head; 1.
DR	Pfam: PR00193; MYOSINHEAVY.
DR	PRINTS: PR00452; SH3DOMAIN.
DR	ProDom: PD000066; SH3; 1.
DR	ProDom: PD000355; myosin_head; 1.
DR	SMART: SM00242; MYSC; 1.
DR	SMART: SM00326; SH3; 1.
DR	PROSITE: PS00002; SH3; 1.
KW	Myosin; Actin-binding; ATP-binding; Chemotaxis; SH3 domain; Multigene family.
KW	
FT	DOMAIN 1 680 MYOSIN HEAD-LIKE.
FT	DOMAIN 681 961 TAIL HOMOLOG REGION 1 (TH.1).
FT	DOMAIN 962 1021 SH3.
FT	DOMAIN 1022 1110 GLY/PRO/ALA-RICH (TH.2).
FT	NP_BIND 101 108 ATP (BY SIMILARITY).
FT	DOMAIN 1016 1113 ALA/GLY/PRO-RICH.
SO	SEQUENCE 113 AA; 124637 MW; 07FEB6C7FA7C2A08 CRC64;



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 106.37 Seconds

(without alignments)
2022.314 Million cell updates/sec

Title: US-09-893-371-1

Sequence: 1 MRYRASALGSDGVRYTMSA.....LITKAKNGHLAVAPRLNSR 1044

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP-archaea:*

2: SP-bacteria:*

3: SP-fungi:*

4: SP-human:*

5: SP-invertebrate:*

6: SP-mammal:*

7: SP-mhc:*

8: SP-organella:*

9: SP-phage:*

10: SP-plant:*

11: SP-rodent:*

12: SP-virus:*

13: SP-vertebrate:*

14: SP-unclassified:*

15: SP-rv1rus:*

16: SP-bacteriaph:*

17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5398	100.0	1044	11 Q9ERB6	Q9ERB6 mus musculus
2	5261	97.5	1028	11 Q63355	Q63355 rattus norv
3	5128	95.0	1028	6 Q27966	Q27966 bos taurus
4	5111	94.7	1028	6 Q28138	Q28138 bos taurus
5	4371	81.0	1028	13 Q92002	Q92002 rana catesb
6	2629	48.7	1035	5 Q87008	Q87008 drosophila
7	1865	34.5	1006	11 Q63357	Q63357 rattus norv
8	1756	32.5	1017	5 Q27328	Q27328 caenorhabdi
9	1719.5	31.9	1099	13 Q90748	Q90748 gallus gall
10	1715	31.8	1109	4 Q12965	Q12965 homo sapien
11	1714.5	31.8	1098	4 Q8MMN7	Q8MMN7 homo sapien
12	1712.5	31.7	1107	11 Q63356	Q63356 rattus norv
13	1595	31.4	1053	5 Q9NGE7	Q9NGE7 strongyloce
14	1573.5	29.1	1186	5 Q61080	Q61080 acanthameob
15	1564	29.0	1100	5 Q19901	Q19901 caenorhabdi
16	1545	28.6	1249	3 Q00647	Q00647 emericella

17	1509.5	28.0	1217	3 Q9Y728	Q9Y728 schizosacch
18	1506	27.9	1049	5 Q00851	Q00851 entamoeba h
19	1504	27.9	633	4 Q96R16	Q96R16 homo sapien
20	1501	27.8	633	4 Q96R15	Q96R15 homo sapien
21	1473	27.3	1215	5 Q77202	Q77202 acanthameob
22	1405	26.0	1036	5 Q17382	Q17382 caenorhabdi
23	1382.5	25.6	1446	10 Q9SWY9	Q9SWY9 arabidopsis
24	1362	25.2	2167	5 Q9V326	Q9V326 drosophila
25	1361	25.2	2357	5 Q9UIM8	Q9UIM8 dicystella
26	1355	25.1	1505	10 Q9LKB9	Q9LKB9 arabidopsis
27	1352	25.0	1515	10 Q93158	Q93158 arabidopsis
28	1346	24.9	839	10 Q9M0G3	Q9M0G3 arabidopsis
29	1336.5	24.8	2179	13 Q9DCG8	Q9DCG8 brachydanio
30	1336	24.7	1528	10 Q24518	Q24518 helianthus
31	1335	24.7	2121	5 Q9NH54	Q9NH54 drosophila
32	1333.5	24.7	2179	13 Q9DCG9	Q9DCG9 brachydanio
33	1331	24.7	1477	10 Q64491	Q64491 arabidopsis
34	1331	24.7	2062	11 Q9JTY5	Q9JTY5 mus musculu
35	1330.5	24.6	2113	11 Q99M26	Q99M26 mus musculu
36	1328	24.6	1547	10 Q8W312	Q8W312 oryza sativ
37	1328	24.6	1611	10 Q8RYE8	Q8RYE8 arabidopsis
38	1328	24.6	2129	5 Q9VLE3	Q9VLE3 drosophila
39	1326	24.6	1260	10 Q24516	Q24516 helianthus
40	1314	24.3	1242	10 Q9M2K0	Q9M2K0 arabidopsis
41	1313.5	24.3	1511	10 Q9M5A6	Q9M5A6 vallneria
42	1312	24.3	1490	10 Q9SKB0	Q9SKB0 arabidopsis
43	1311.5	24.3	1520	10 Q39160	Q39160 arabidopsis
44	1311	24.3	1515	10 Q9ATB7	Q9ATB7 petroselinu
45	1310.5	24.3	1529	10 Q9XE14	Q9XE14 zea mays (m

ALIGNMENTS

RESULT 1

ID Q9ERB6 PRELIMINARY; PRT; 1044 AA.

AC Q9ERB6;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Nuclear myosin I beta.

GN MYOIC.

OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RM [1]

RP SEQUENCE FROM N.A.

RA Peatic-Dragovich L., Stojiljkovic L., Phillimonenko A.A., Nowak G.,

RA Ke Y., Settleage R.E., Shabanowitz J., Hunt D.F., Hozak P.,

RA de lanerolle P.;

RT "A Myosin I Isoform in the Nucleus.";

RL Science 0:0-0(2000).

DR EMBL; AY007255; AAC02570.1; -.

DR HSSP; P08799; IMND.

DR MGD; MGI:106612; MYOIC.

DR InterPro; IPR000048; IO_region.

DR InterPro; IPR001609; myosin_head.

DR Pfam; PF00612; IQ; 3.

DR Pfam; PF00663; myosin_head; 1.

DR PRINTS; PR00193; MYOSINHEAVY.

DR	ProDom; PD000355; myosin_head; 1.
DR	SMART; SM00015; IQ; 2.
DR	SMART; SM00242; MYSC; 1.
SO	SEQUENCE 1044 AA; 119876 MW; CDAB0FF699D041C9 CXC64;

Query Match 100.0%; Score 5398; DB 11; Length 1044;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRYTMSA...LITKAKNGHLAVAPRLNSR 1044

```

Db      1  MRYRASNLSGSDVRYTMESEALTRADRVGVODFVLENTSEAFIENLRFRRENLIYTY 60
QY      61  IGVPLVSNPFYRDLQIYSRQHMERYGVSEYEPHLEFVADTVYRALFERRDQAVMTS 120
Db      61  IGVPLVSNPFYRDLQIYSRQHMERYGVSEYEPHLEFVADTVYRALFERRDQAVMTS 120
QY      121  GESGAGKTEATKRLLOFYAETCPAPERGAVERDRLLOSNPVELEAFGNATLRNDNSRRG 180
Db      121  GESGAGKTEATKRLLOFYAETCPAPERGAVERDRLLOSNPVELEAFGNATLRNDNSRRG 180
QY      181  KYMDVDFEFGAVGVGHILSYLLEKSRVYHONGERNHFVYOLLGGEETLRRLGLER 240
Db      181  KYMDVDFEFGAVGVGHILSYLLEKSRVYHONGERNHFVYOLLGGEETLRRLGLER 240
QY      241  NPOSYLYLVKGQCAKAVSINDSKDMVKRKALSYIDFTEDEVEDLSTIVASVHLGNHIF 300
Db      241  NPOSYLYLVKGQCAKAVSINDSKDMVKRKALSYIDFTEDEVEDLSTIVASVHLGNHIF 300
QY      301  AAEDESNMAOYTENOLKYTLRILGVGGTTLREALFTRKIIAKGELLSPNLBQAAYARD 360
Db      301  AAEDESNMAOYTENOLKYTLRILGVGGTTLREALFTRKIIAKGELLSPNLBQAAYARD 360
QY      361  ALAKAVYSRTFTWLVYRKINSLSKDAESPMSRSTVGLDLYGFEVQHNSEQFCIN 420
Db      361  ALAKAVYSRTFTWLVYRKINSLSKDAESPMSRSTVGLDLYGFEVQHNSEQFCIN 420
QY      421  YCNEKIQOOLFTELTLKSEOEYEAEGIAMPEVQYFNKKIICDIVEEKFGIISILDEEC 480
Db      421  YCNEKIQOOLFTELTLKSEOEYEAEGIAMPEVQYFNKKIICDIVEEKFGIISILDEEC 480
QY      481  RPEADTDLFLEKLEEDVVKPHPHFLTHKLADQKTRSLDGEFRLHYAGEVYVSYTGEF 540
Db      481  RPEADTDLFLEKLEEDVVKPHPHFLTHKLADQKTRSLDGEFRLHYAGEVYVSYTGEF 540
QY      541  DKNNDLFRNLKETTMOSSNMPIMAOCFKSELSDDKRPETVATQFKMSLQIYVEIIRSE 600
Db      541  DKNNDLFRNLKETTMOSSNMPIMAOCFKSELSDDKRPETVATQFKMSLQIYVEIIRSE 600
QY      601  PAYIRCTKPRDAKOPGFEDVLIRHOKYVIGLMENIRVRAGAVYRKKEAFIQRKSKC 660
Db      601  PAYIRCTKPRDAKOPGFEDVLIRHOKYVIGLMENIRVRAGAVYRKKEAFIQRKSKC 660
QY      661  PETWPMWAGRPDGVAVLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLLEVROSLAT 720
Db      661  PETWPMWAGRPDGVAVLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLLEVROSLAT 720
QY      721  KIQAAWGFHWKQKFLYVKSALCISQSWMRGTIGRRKAARKMAAOTIRRLIGFILRHS 780
Db      721  KIQAAWGFHWKQKFLYVKSALCISQSWMRGTIGRRKAARKMAAOTIRRLIGFILRHS 780
QY      781  PRCEENAFIDHVRASFLNLRLRQLPNNVLDTSMPTRPPALREASELRLCLCKNNVMYK 840
Db      781  PRCEENAFIDHVRASFLNLRLRQLPNNVLDTSMPTRPPALREASELRLCLCKNNVMYK 840
QY      841  CRISISEPMKQOOLQKAVASEIFKGGKDNYPQSVPRLFISRLGTEESIPRVLQSLGSEPT 900
Db      841  CRISISEPMKQOOLQKAVASEIFKGGKDNYPQSVPRLFISRLGTEESIPRVLQSLGSEPT 900
QY      901  QYAVPVYKDYRKCYKPPRQOLLTPSAVIVVEDAKKQRIIDVANTLGISVSSLSDFLV 960
Db      901  QYAVPVYKDYRKCYKPPRQOLLTPSAVIVVEDAKKQRIIDVANTLGISVSSLSDFLV 960
QY      961  HVQEDNRKQGVVLOSQDHYIETLTALTSADRVNININQGSITPAGGGRGIIIDFTS 1020
Db      961  HVQEDNRKQGVVLOSQDHYIETLTALTSADRVNININQGSITPAGGGRGIIIDFTS 1020
QY      1021  GSELLITKAKNGHLAVVAPRLNSR 1044
Db      1021  GSELLITKAKNGHLAVVAPRLNSR 1044

```

RESULT 2
063355

```

ID      063355  PRELIMINARY:      PRT: 1028 AA.
AC      063355:
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE      Myosin I heavy chain.
GN      MYR 2.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPRAGUE-DAWLEY;
RA      Ruppert C., Godel J., Reinhard J., Brehler M.;
RT      "Myr-2, a novel class-I myosin identified in rat brain.";
RL      Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
DR      EMBL; X74800; CAA52807.1; -.
DR      HSSP; P08799; 1MND.
DR      InterPro: IPR000048; IQ_region.
DR      InterPro: IPR001609; myosin_head.
DR      Pfam; PF00612; IQ_3.
DR      Pfam; PF00063; myosin_head; 1.
DR      PRINTS; PR00193; MYOSINHEAVY.
DR      ProDom; PD000355; MYOSINHEAVY.
DR      SMART; SM00015; IQ; 2.
DR      SMART; SM00242; MYSC; 1.
SQ      SEQUENCE 1028 AA: 118090 MW: 10E3B1AFD5999CB5 CRC64;

Query Match      97.5%; Score 5261; DB 11; Length 1028;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1015; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      17  MESALTARDRVGVODFVLENTSEAFIENLRFRRENLIYTYIGVPLVSNPFYRDLQI 76
Db      1  MESALTARDRVGVODFVLENTSEAFIENLRFRRENLIYTYIGVPLVSNPFYRDLQI 60
QY      77  YSRQHMERYGVSEYEPHLEFVADTVYRALFERRDQAVMISGSGAGKTEATKRLLO 136
Db      61  YSRQHMERYGVSEYEPHLEFVADTVYRALFERRDQAVMISGSGAGKTEATKRLLO 120
QY      137  FYAETCPAPERGAVERDRLLOSNPVELEAFGNATLRNDNSRRGIIIDFTS 196
Db      121  FYAETCPAPERGAVERDRLLOSNPVELEAFGNATLRNDNSRRGIIIDFTS 180
QY      197  HITSYLLEKSRVYHONGERNHFVYOLLGGEETLRRLGLERNPQSYLYLVKGQCAV 256
Db      181  HITSYLLEKSRVYHONGERNHFVYOLLGGEETLRRLGLERNPQSYLYLVKGQCAV 240
QY      257  SSINDSKDMVKRKALSYIDFTEDEVEDLSTIVASVHLGNHIFAEDESNMAOYTENOL 316
Db      241  SSINDSKDMVKRKALSYIDFTEDEVEDLSTIVASVHLGNHIFAEDESNMAOYTENOL 300
QY      317  KYTLRLGVGGTTLREALFTRKIIAKGELLSPNLBQAAYAROLAAYVSRFTWLVYR 376
Db      301  KYTLRLGVGGTTLREALFTRKIIAKGELLSPNLBQAAYAROLAAYVSRFTWLVYR 360
QY      377  KINRSLASKDAESPMSRSTVGLDLYGFEVQHNSEQFCINCNKELQOLFTELTLK 436
Db      361  KINRSLASKDAESPMSRSTVGLDLYGFEVQHNSEQFCINCNKELQOLFTELTLK 420
QY      437  SEOEYEAEGIAMPEVQYFNKKIICDIVEEKFGIISILDEECILPAGATDLFLEKED 496
Db      421  SEOEYEAEGIAMPEVQYFNKKIICDIVEEKFGIISILDEECILPAGATDLFLEKED 480
QY      497  TYKPRPHFLTHKLAOKTRKSLDGEFRLHYAGEVYVSYTGEFLDKNNDLFRNLKETMC 556
Db      481  TYKPRPHFLTHKLAOKTRKSLDGEFRLHYAGEVYVSYTGEFLDKNNDLFRNLKETMC 540
QY      557  SSNMPIMAOCFKSELSDDKRPETVATQFKMSLQIYVEIIRSEKPEAYTRCIKPDNAKOPG 616
Db      541  SSNMPIMAOCFKSELSDDKRPETVATQFKMSLQIYVEIIRSEKPEAYTRCIKPDNAKOPG 600

```



```

Oy 617 REDEVLIIRHGVYKYLGLMENTLRYVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPQDGYA 676
    |||||||
Db 601 RFDEVLIIRHGVYKYLGLMENTLRYVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPQDGYA 660
Oy 677 VLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRROSLATKIQAMRGFHWKQKFL 736
    |||||||
Db 661 VLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRROSLATKIQAMRGFHWKQKFL 720
Oy 737 RKRSATCICQSMWRGTGRRKAKRMAAOTIRRLIRGFTLRHSRCPENAFELDHVYAS 796
    |||||||
Db 721 RKRSATCICQSMWRGTGRRKAKRMAAOTIRRLIRGFTLRHSRCPENAFELDHVYAS 780
Oy 797 FLNLRLROLPRNVLDTSWPTPPALREASSELLRELCMKNMVKYCRSISPENKQOLOKA 856
    |||||||
Db 781 FLNLRLROLPRNVLDTSWPTPPALREASSELLRELCMKNMVKYCRSISPENKQOLOKA 840
Oy 857 VASEIFKGGKNDYPOSVPRLEFISTRUGTEESIPRYLQSLGSEPIQYAVPVYKYDRKGYP 916
    |||||||
Db 841 VASEIFKGGKNDYPOSVPRLEFISTRUGTEESIPRYLQSLGSEPIQYAVPVYKYDRKGYP 900
Oy 917 RPRQLLTTPSAVYIVEDAKVKORIDYANLGTISVSSLSDSL FVLHVOREDNKKOKGDVYLQ 976
    |||||||
Db 901 RSRQLLTTPSAVYIVEDAKVKORIDYANLGTISVSSLSDSL FVLHVOREDNKKOKGDVYLQ 960
Oy 977 SDHVIETLTKTALSDRVNNININGSIITFAGGPGRGDIIDFTSGSELLITKANGLHAY 1036
    |||||||
Db 961 SDHVIETLTKTALSDRVNNININGSIITFAGGPGRGDIIDFTSGSELLITKANGLHAY 1020
Oy 1037 VAPRLNSR 1044
    |||||||
Db 1021 VAPRLNSR 1028

RESULT 3
027966 PRELIMINARY; PRT; 1028 AA.
ID 027966
AC 027966:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Myosin I.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADRENAL GLAND;
RX MEDLINE=94148088; PubMed=8313976;
RA Zhu T., Ikebe M.;
RT "Cloning of myosin I from bovine adrenal gland.";
RL FEBS Lett. 339:31-36(1994).
DR EMBL; U03420; AAA17565.1; -.
DR HSSP; P08799; IAMD.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head.1.
DR PRINTS; PR00193; MYOSINHEAD.
DR PRODOM; PD000353; MYOSINHEAD.1.
DR SMART; SM00015; IQ; 2.
DR SMART; SM00242; MYSC; 1.
SQ SEQUENCE 1028 AA; 118020 MW; 5CFEE28CF2A52D7 CRC64;

```

Query Match 95.0%; Score 5128; DB 6; Length 1028;
 Best Local Similarity 95.6%; Pred. No. 0;
 Matches 983; Conservative 25; Mismatches 20; Indels 0; Gaps 0;

```

Oy 17 MESALTARDRVGVDFVLENTSEAAFIENLRRFRNLTYYTIGPYLVSNPYRDLQI 76
    |||||||
Db 1 MESALTARDRVGVDFVLENTSEAAFIENLRRFRNLTYYTIGPYLVSNPYRDLQI 60

```

```

Oy 77 YSRQHMERGVGFSEYEVPHLFAVADTVYRALPERRDOAVMTSGESAGKTEATRLQ 136
    |||||||
Db 61 YSRQHMERGVGFSEYEVPHLFAVADTVYRALPERRDOAVMTSGESAGKTEATRLQ 120
Oy 137 FYAETCPAPERGAVRDRLLOSNPVLEAFGNATKLRLDNNSSRGKYADVDFDKGAPVG 196
    |||||||
Db 121 FYAETCPAPERGAVRDRLLOSNPVLEAFGNATKLRLDNNSSRGKYADVDFDKGAPVG 180
Oy 197 HILSYLLEKSRVYQNHNGERNFHVQVQLGEGEETRLRLGLEBNPOSTYLKVGQCAKY 256
    |||||||
Db 181 HILSYLLEKSRVYQNHNGERNFHVQVQLGEGEETRLRLGLEBNPOSTYLKVGQCAKY 240
Oy 257 SSINDKDMKVMKRALVIDFTEDEVEDLSIYASVYLHGNHFPADDESNADVTENOL 316
    |||||||
Db 241 SSINDKDMKVMKRALVIDFTEDEVEDLSIYASVYLHGNHFPADDESNADVTENOL 300
Oy 317 KYLTRLGVEGTTLREALTHRKIIJAKGELLSPNLBOAAVARDALAKAVYSTFTWLVR 376
    |||||||
Db 301 KYLTRLGVEGTTLREALTHRKIIJAKGELLSPNLBOAAVARDALAKAVYSTFTWLVR 360
Oy 377 KINRSLASKDAESFSMSTVVLGLDLYGGEVQHNSEFEOPCINCEKIQOLFIELTLK 436
    |||||||
Db 361 KINRSLASKDAESFSMSTVVLGLDLYGGEVQHNSEFEOPCINCEKIQOLFIELTLK 420
Oy 437 SEQEYEAEGIAWEPVOYFNKKIICDLYEEKFKGIISILDEECLRPEATDLFLEKLED 496
    |||||||
Db 421 SEQEYEAEGIAWEPVOYFNKKIICDLYEEKFKGIISILDEECLRPEATDLFLEKLED 480
Oy 497 TVKPHPHFLTHKLADQTKRSLDGEFFRLHYAGEVTVYTGFLDKNNDLFFNLKETWC 556
    |||||||
Db 481 TVKPHPHFLTHKLADQTKRSLDGEFFRLHYAGEVTVYTGFLDKNNDLFFNLKETWC 540
Oy 557 SSNMPIMAOCPDSESDKKRPETVATQEFMSLQOLVETLRKSEPAYIRIKRNDAKOPG 616
    |||||||
Db 541 SSNMPIMAOCPDSESDKKRPETVATQEFMSLQOLVETLRKSEPAYIRIKRNDAKOPG 600
Oy 617 RFDEVLIIRHGVYKYLGLMENTLRYVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPQDGYA 676
    |||||||
Db 601 RFDEVLIIRHGVYKYLGLMENTLRYVRAGFAVRRKYEAFLORYKSLCPETWPMAGRPQDGYA 660
Oy 677 VLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRROSLATKIQAMRGFHWKQKFL 736
    |||||||
Db 661 VLVNHLGYKPEEYKMGRTKIFIRPKTLFATEDSLEVRROSLATKIQAMRGFHWKQKFL 720
Oy 737 RKRSATCICQSMWRGTGRRKAKRMAAOTIRRLIRGFTLRHSRCPENAFELDHVYAS 796
    |||||||
Db 721 RKRSATCICQSMWRGTGRRKAKRMAAOTIRRLIRGFTLRHSRCPENAFELDHVYAS 780
Oy 797 FLNLRLROLPRNVLDTSWPTPPALREASSELLRELCMKNMVKYCRSISPENKQOLOKA 856
    |||||||
Db 781 FLNLRLROLPRNVLDTSWPTPPALREASSELLRELCMKNMVKYCRSISPENKQOLOKA 840
Oy 857 VASEIFKGGKNDYPOSVPRLEFISTRUGTEESIPRYLQSLGSEPIQYAVPVYKYDRKGYP 916
    |||||||
Db 841 VASEIFKGGKNDYPOSVPRLEFISTRUGTEESIPRYLQSLGSEPIQYAVPVYKYDRKGYP 900
Oy 917 RPRQLLTTPSAVYIVEDAKVKORIDYANLGTISVSSLSDSL FVLHVOREDNKKOKGDVYLQ 976
    |||||||
Db 901 RSRQLLTTPSAVYIVEDAKVKORIDYANLGTISVSSLSDSL FVLHVOREDNKKOKGDVYLQ 960
Oy 977 SDHVIETLTKTALSDRVNNININGSIITFAGGPGRGDIIDFTSGSELLITKANGLHAY 1036
    |||||||
Db 961 SDHVIETLTKTALSDRVNNININGSIITFAGGPGRGDIIDFTSGSELLITKANGLHAY 1020
Oy 1037 VAPRLNSR 1044
    |||||||
Db 1021 VAPRLNSR 1028

```

RESULT 4
 028138 PRELIMINARY; PRT; 1028 AA.
 ID 028138
 AC 028138;

DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Myosin IB.
 OS Bos taurus (Bovinae).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NC NCBL_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRNIN;
 RX MEDLINE=94294379; PubMed=8022785;
 RA Reitzes O., Barylko B., Li C., Sudhof T.C., Alberici J.P.;
 RT "Domain structure of a mammalian myosin I-beta";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6349-6353(1994).
 DR EMBL: 222852; CA80476.1; -.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAV.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 2.
 DR SMART: SM00242; MYSC; 1.
 SQ SEQUENCE 1028 AA; 118241 MW; 95E7E806CC6B2CD CRC64;

Query Match 94.7%; Score 5111; DB 6; Length 1028;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 980; Conservative 27; Mismatches 21; Indels 0; Gaps 0;

QY 17 MESALTRAPRGVQVDVLENTSEAFTEMLRRRREMLITYTGPVAVSNPYDLOI 76
 DB 1 MESALTRAPRGVQVDVLENTSEAFTEMLRRRREMLITYTGPVAVSNPYDLOI 60
 QY 77 YSHQHMERKGVSEFYVPPHFAVADTVYRALTERERDAVMISGSGAGKTEATKRLQ 136
 DB 61 YSHQHMERKGVSEFYVPPHFAVADTVYRALTERERDAVMISGSGAGKTEATKRLQ 120
 QY 137 FYAETCPAPRGVAVDRLLQSNPVLFAFGNAKTLENDSSRRGKYMVDQFPGAPVGG 196
 DB 121 FYAETCPAPRGVAVDRLLQSNPVLFAFGNAKTLENDSSRRGKYMVDQFPGAPVGG 180
 QY 197 HILSYLLEKSRVYHONHGERNFHVEYQLGEGEETRLGLERNOSLYLYKGGCAV 256
 DB 181 HILSYLLEKSRVYHONHGERNFHVEYQLGEGEETRLGLERNOSLYLYKGGCAV 240
 QY 257 SSINDKSDKVMKALSVIDFTEDEVEDLSTIYASVLAHGNTHFADEDSNAQVTEENQ 316
 DB 241 SSINDKSDKVMKALSVIDFTEDEVEDLSTIYASVLAHGNTHFADEDSNAQVTEENQ 300
 QY 317 KYTRLRGVGTTLREALTHRTKIAKGEELSPLEEQAYARDALAKAVYSRTFTWLYR 376
 DB 301 KYTRLRGVGTTLREALTHRTKIAKGEELSPLEEQAYARDALAKAVYSRTFTWLYR 360
 QY 377 KIRSLASKDAESPMSRSTTVGLDIYGFVEYQHSFQFCNTCNENKLOQFIELTLK 436
 DB 361 KIRSLASKDAESPMSRSTTVGLDIYGFVEYQHSFQFCNTCNENKLOQFIELTLK 420
 QY 437 SEOEYEAEGIAPEVOYFNKILCDLVEKEKGIISILDECLRGEATDLEFLEKLD 496
 DB 421 SEOEYEAEGIAPEVOYFNKILCDLVEKEKGIISILDECLRGEATDLEFLEKLD 480
 QY 497 TYKPHFLTHKTLADOKTRKSLDRGFRLHYAGVTVSYTGFLKNNDLRLNKEIYC 556
 DB 481 TYKPHFLTHKTLADOKTRKSLDRGFRLHYAGVTVSYTGFLKNNDLRLNKEIYC 540
 QY 557 SSNNPIMAOOFKSEISDKKRPETVTOFKMSILOVELIRSEKAPYIRICIPNDKOG 616
 DB 541 SSNNPIMAOOFKSEISDKKRPETVTOFKMSILOVELIRSEKAPYIRICIPNDKOG 600
 QY 617 RFDEVLIHQVYKYLGIEMNLVRRAGFAVRKRYEAFLOQYKSLCPETPMWAGRPDGYA 676

DB 601 RFDEVLIHQVYKYLGIEMNLVRRAGFAVRKRYEAFLOQYKSLCPETPMWAGRPDGYA 660
 QY 677 VLVRLHGYKPEEYKMGRTKIFIRPEKTLFATEDSLEVRQSLATKIOAMRGFWHROKFL 736
 DB 661 VLVRLHGYKPEEYKMGRTKIFIRPEKTLFATEDSLEVRQSLATKIOAMRGFWHROKFL 720
 QY 737 RVKRSALICISWMWGTGRRKRAKRWMAQITRLNGFLRHSPROENAFDHYRAS 796
 DB 721 RVKRSALICISWMWGTGRRKRAKRWMAQITRLNGFLRHSPROENAFDHYRAS 780
 QY 797 FLNLRQPLPNNVLDISWPTPPALREASELLRELCNMWYKCRSISPEMKOOLQOKA 856
 DB 781 FLNLRQPLPNNVLDISWPTPPALREASELLRELCNMWYKCRSISPEMKOOLQOKA 840
 QY 857 VASEIFKGNKNDYQVPRLEFISITRLGTEISPYVLOSLEPQIYAVPYKIDRKGYKP 916
 DB 841 VASEIFKGNKNDYQVPRLEFISITRLGTEISPYVLOSLEPQIYAVPYKIDRKGYKP 900
 QY 917 RPROLLTPSAVVTVEKAKVORIDYANLIGISVSLSDSLFVLYHOREDKOKGDPVLO 976
 DB 901 RPROLLTPSAVVTVEKAKVORIDYANLIGISVSLSDSLFVLYHOREDKOKGDPVLO 960
 QY 977 SDHVLETLTKTALSADRVNNININGSLTFAGGGRDGIIDFTSGSELLTKAKNGHLAV 1036
 DB 961 SDHVLETLTKTALSADRVNNININGSLTFAGGGRDGIIDFTSGSELLTKAKNGHLAV 1020
 QY 1037 VAPRLNSR 1044
 DB 1021 VAPRLNSR 1028

RESULT 5

ID 092002 PRELIMINARY; PRT: 1028 AA.
 AC 092002;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Myosin I beta.
 GN AMT BETA.
 OS Rana catesbeiana (Bull. frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
 NC NCBL_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=95083594; PubMed=7991542;
 RA Metcalf A.B., Chelliah Y., Hudspeeth A.J.;
 RT "Molecular cloning of a myosin I beta isozyyme that may mediate
 adaptation by hair cells of the bullfrog's internal ear";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11821-11825(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Solc C.F., Deflier B.H., Duyk G.M., Corey D.P.;
 RL Aud. Neurosci. 1:63-75(1994).
 DR EMBL: U14549; AA57192.1; -.
 DR EMBL: U14382; AA65091.1; -.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IO_region.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR PRINTS: PR00193; MYOSINHEAV.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 2.
 DR SMART: SM00242; MYSC; 1.
 SQ SEQUENCE 1028 AA; 118830 MW; D15F6F99B86A90AE CRC64;

Query Match 81.0%; Score 4371; DB 13; Length 1028;
 Best Local Similarity 79.2%; Pred. No. 1; 2e-292;


```

Db      597 KPNDDQJANVFNDYVLVLHGVKYLGLMENVRRAGFAFRYTELEERKSKSTKSTWPNV 656
Qy      668 --AGRPQDGVAVLVRLHGYKPEEYKMGRTKIFIRPEKTLFATEDSLVAVROSLATKIOAA 725
Db      657 KRPGRKAGVQOVLVDLGDWDEKRYNGEKLEIRNPRTLFEDVAVQEKHEIAALIQAH 716
Qy      726 WRFQFMRKQFLNFKNSAICIGSMWRGTGRRKRAKRWAAQTIRLIRGFIIRHSPRDE 785
Db      717 WKGIMORRRKYLKRAQVILIMOSYCRKLAQAARKREAAADKIRAFIKGITRRNDAPNGF 776
Qy      786 NAFELDHVAVASFLNLRQLPNNVDTQSPTRPPALREASSELLRELCMKNNWVKCRSIS 845
Db      777 NEEFLINAKRMMLRLAKELPTKVLDDKSPHAPGCEESGLHRLHRLHRLKIRKLIT 836
Qy      846 PEMKQLOQKAAVASELFGKKNYPOSVRLFISTRLGTEET--SPRYLQSLGSEPIQY 902
Db      837 POKRQFELKVLAEKVEKGNKNVYASVSTWFOEDRIPKEHIOVNDVFASTFGSEQLKY 896
Qy      903 AYPVAVYDKKGYKPRRROLLTPSAVYVEDAKVQ-----RIDYALTGISVSSLS 955
Db      897 QSECTKFDHNGYKSRDFILSNKALYVLGKTYQKHLPLDKIDF-----TLTNND 950
Qy      956 SLFVLHVOREDNKQKGDVYVLOSQSDHVEITLKTALSDRVNNINI--NQSITFAGSPGRD 1014
Db      951 DLNVRIRPLDKKDKGDLILIPRIEFSTYIIDYGTASIVSYVDRNLHNHYVKGKG 1010
Qy      1015 IIDFTSGSELLITKAKNGLAVV 1037
Db      1011 VIDTGTGAEPGVVRDK--GHLVIT 1032

```

```

RESULT 7
ID 063357 PRELIMINARY; PRT: 1006 AA.
AC 063357;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Myosin I.
GN MYR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=94308268; PubMed=8034741;
RA Baehler M., Kroschewski R., Stoeffler H.E., Behrmann T.;
  "Rat myr4 defines a novel subclass of myosin I: Identification,
  RT distribution, localization, and mapping of calmodulin-binding sites
  with differential calcium sensitivity."
RT J. Cell Biol. 126:375-389(1994).
RL EMBL: X71997; CAAS0871.1; -.
DR HSSP: P08799; IAMD.
DR InterPro: IPR000048; IO_region.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IO_2.
DR Pfam: PF00063; myosin_head; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR Prodom: PD000355; myosin_head; 1.
DR SMART: SM00015; IO_1.
DR SMART: SM00242; MYSC; 1.
SQ SEQUENCE 1006 AA; 116137 MW; 9AA0B626A0FDA42A CRC64;

```

Query Match 34.5%; Score 1865; DB 11; Length 1006;
 Best Local Similarity 41.7%; Pred. No. 8.5e-120;
 Matches 426; Conservative 158; Mismatches 301; Indels 136; Gaps 26;

```

Qy      28 GVQDFVLENTFSEAFIENLRRFRRENLIYTYIGVLYSVNRYRLOQIYSRQHMERYG 87
Db      10 GKADFVLMDF--VSMPEFMANLRLRFKGRITYPEIGEVVSVNPKYVNLITYGROTIQYKG 68

```

```

Qy      88 VSEYEVPHFLFAVADIVYRALTRERDQAVMISGSEAGKTEATKRLLOFAETCAPER 147
Db      69 RELPERFLPAIIDAIVYKAKMRKSKPTCLMISSESGAGTEVSKTYMOTIAITNPSOR 128
Qy      148 G--GAVRDLQSNPVLAEFGNAKTLRNDSNRGKYMDVQDFPKAPVGHILSYLLEK 205
Db      129 AEIERVKMMLKSCVLEAFPMNAKTNRNDSNRGKYMDVQDFPKAPVGHILSYLLEK 188
Qy      206 SRVYHONHGERNFVYTOLEEGEETLRRLGERNQSTLYLYKGCAVSSINDKSDW 265
Db      189 SRVYVQGGERSFSFYOQLQGGSEOMHSJHLQKSSSYNYLVGAGQL--SSINDAAEF 247
Qy      266 KVMKALSIDFTDEVEDLSTVASYLHGNIFPADEDS---NAQVTENQKLYTR 321
Db      248 KVVADAMVYGFKEPEIDYTKILAILHLGNLFYDGDPTLENGK-----VSVIAE 302
Qy      322 LLGVEGTLRLREALTRKLIANGEBLLSPLEQ--AAVADALAKAVYSRFTWLVRKLR 380
Db      303 LLSTRKADWVERALLYR--TVATGRDIDIKQHTGEQASVGRDAFAKAIYERLFCWIVTRIND 361
Qy      381 SLASDAESPBRSTYVGLDIYGFVEYFOHNSFEQICINCNKLOQLFTELTKSEOE 440
Db      362 IIEKKNYDTVHGKNTYIGVLDIYGFELFDNNSFEQICINCNKLOQLFQVLKQOE 421
Qy      441 EYEAEGIAPEVQYFNKNIIDIVEEKKGIISLDECLRPGBATDLPTELKLEDTYKP 500
Db      422 EYQREGIPWKHIDFNNQIIVDVEQOKGIALLDQACNVGAVTGCMFLEALNSKLG 481
Qy      501 HPHELYTHKLADQKTRKSLDRGEFRLIYAGVYTVSVTGLDKNDLLPRNLKEMCSMN 560
Db      482 HGHESSRKTCASDKLIEFDR--DFRIRYHAGVYTVSVTGLDKNDLTFQDKRLMYSSN 540
Qy      561 PIMACPEKSELS--DKRPETATOPKMSILOVETILRSEPAVIRICIPNDAKOPGR 617
Db      541 PVLKMMPEGKLISTEVTKRPPLTAITLFKSNMIALVNDLASKETYYRCAIKPNKKSPQI 600
Qy      618 FDEVLIHQVYGLIMENLRVRAGFAVRRRYEAFLOKYSLCEPTWMMAGRPQDGVAV 677
Db      601 FDDERCQHQVEYGLLENVVRBRAGFAVRYTEKFLHRYKMSISETWPN--HDLPSDEAV 659
Qy      678 --LVNHLGKPEEYKMGRTKIFIRPEKTLFATEDSLVAVROSLATKIOAAMRGHMKOF 735
Db      660 KKLIRRCFO--DVAAYGKTKIFIRPRLFLE--ELRAQMLV----- 699
Qy      736 LKVRSAICIGSMWRGTGRRKRAKRWAAQTIRLIRGFIIRHSPRCPENAFPLDHVRA 795
Db      700 ----RVVLFQKVMRGITLARMR--YKRTAALITIRYRYRYK----- 737
Qy      796 SFNLNLRQLP--PRNVLD-----TSWPTPPALREASSELLRELCMKNNWVKCRSISPEWK 849
Db      738 SYIHVARRFRHGVKNMRYGKHVKPPTPPKYLRRFEELQSIIFNRASOLIKTIIPASDL 797
Qy      850 QOLOKAAVASELFGKKNYPOSVRLFISTRLGTEETSPVLOSISGSEPIQYAVPV-- 907
Db      798 PQVRAKVAAMELTKQRADL--GLORAMEGNYLASKPPTP---OTSGL-----FVPAANE 847
Qy      908 ----KY-----DRKGYKPRROLLTPSAVYVEDAKV 936
Db      848 LKRRKKNMNVLESCHVRKVRNRSKVEDAIFVTBRLHLYKMP-----TKQYK 895
Qy      937 KORIDYANLITGISVSSLSLFLVHVOREDNKQKGDVYVLOSQSDHVEITLKTALSDRVNN 996
Db      896 KMTIPLYNLTGISVNGRDOJLVFHT--KDNK-----DLVCLFSKOPTHESTRIGE 944
Qy      997 I 997
Db      945 I 945

```

RESULT 8
 ID 027328 PRELIMINARY; PRT: 1017 AA.

```

AC      Q27328.10V-1996 (TremblRel. 01, Created)
DT      01-NOV-1996 (TremblRel. 01, Last sequence update)
DT      01-JUN-1996 (TremblRel. 01, Last sequence update)
DT      01-JUN-2002 (TremblRel. 21, Last annotation update)
DE      HMW-5 protein.
GN      MCE1A.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC      Rhabditidae; Pelodierinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RL      Lloyd C.;
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Cope M.J.T.V., Kendrick-Jones J.;
RL      Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RL      EMBL; Z35603; CA84673.1; -.
DR      EMBL; X75564; CA53244.1; -.
DR      HSSP; P08799; 1MND.
DR      InterPro; IPR000048; IO_region.
DR      InterPro; IPR001609; myosin_head.
DR      Pfam; PF00612; IQ; 1.
DR      Pfam; PF00063; myosin_head; 1.
DR      PRINTS; PR00193; MYOSINHEAVY.
DR      ProDom; PD000355; myosin_head; 1.
DR      SMART; SM00015; IQ; 1.
DR      SMART; SM00242; MYSC; 1.
KW      Myosin.
SQ      SEQUENCE 1017 AA; 116557 MW; 461FF63A2C955ED8 CRC64;

```

Query Match	32.5%;	Score 1756;	DB 5;	Length 1017;
Best Local Similarity	39.1%;	Pred. No. 2.9e-112;		
Matches 405;	Conservative 179;	Mismatches 329;	Indels 124;	Gaps 24

[illegible][illegible]

RESULT 9
Q90748
ID Q90748 PRELIMINARY;
PRT; 1099 AA

```

DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Brush border myosin IB.
GN      CBM1B.
OC      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=INTESTINE;
RA      Knight A.E., Kendrick-Jones J.;
RT      "A novel vertebrate myosin I."
RL      Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: CONTRAINS 1 SH3 DOMAIN.
DR      EMBL; X70400; CAA49850.1; -.
DR      HSSP; P08799; 1MND.
DR      InterPro; IPR000048; IQ_region.
DR      InterPro; IPR001609; myosin_head.
DR      InterPro; IPR001452; SH3.
DR      Pfam; PF00612; IQ. 1.
DR      Pfam; PF00063; myosin_head; 1.
DR      Pfam; PF00018; SH3. 1.
DR      PRINTS; PR00193; MYOSINHEAVY.
DR      PRINTS; PR00452; SH3DOMAIN.
DR      ProDom; PD000066; SH3. 1.
DR      ProDom; PD000355; myosin_head; 1.
DR      SMART; SM00015; IQ. 1.
DR      SMART; SM00242; MYSC. 1.
DR      SMART; SM00326; SH3. 1.
DR      PROSITE; PS50002; SH3. 1.

```


OY		328	TTLRALHRLKILIA-----KGELLSPNLDEQAAYARDAALAKAVSRTPMLVLRKINRSLA	383
Dd		317	DLRKELTLLSRODSSWMGGSSESIHYTLVNEQACYTRDALAKLARHVFDELVDLSINKAW-	375
OY		364	SKDASPMSRSTVTYGLLDIYGFEEFQHNHSFEQFCINYNCKEQLQFLTELTKSHOEYE	443
Dd		376	EKDHEBYN-----IGVLIDYGFELFGNGNFQPCINYNENKLQFIETLTKAQDEEYV	429
OY		444	AEGIMEVEVOYFNKKITCDLYBEERK--GIISIDECL---RPGEAIDLFFLEKLEDTV	498
Dd		430	QEGIMTPIETEFNNKIYCDLIENKNVPGISMIISLDVCATMHAIVEGAQDTLLQIKLOMDI	489
OY		499	KPHPFELHLKADQDKTRKSIDRGCEPRLHHYGEVYVSYTGFKRNNDLFRLKRETMSS	558
Dd		490	GSHEIF-----NSMWOG-FIIHHYGKYSIDMGCFERRRDVLFMDLIELMOSS	537
OY		559	MNPIMAOCEFDSSELSDK-RPEYVAIQPKMSILLQVELIRSKEPAYICIRKPADAQOPGR	617
Dd		538	ELRFKTSLEPENLQDKKRPTTAGSKIKKQANDLVSTLMCTPHYINCINKEKRRKRD	597
OY		618	FDEVILIRQVNYTGLGMENLRYRRRAGFAVRKKEAFLOQRKSLCPEPTPMMAGRPOGVAV	677
Dd		598	WEESRVKHQIVELGLKENIRVARRAQYARRIFOKFLQRYALTITKATWSWGCEEKQVYLH	657
OY		678	LVRHLGYRPEEKMKMGRTKIFTRFPKTLPATDESIVBRQSATKTIOAMRGPHMRQKFLR	737
Dd		658	LLQSNMMSDDPOLGRSKVFTIKAPESLLEEMERKTDGYARVIQKSMRKTFVARKKYVO	717
OY		738	VKRSHAIQSWMRGTLGRRKAARRWAAOTIRRLRGFILRHSPRCPENAEFLDHVASF	797
Dd		718	MR-----	719
OY		798	LLNLRLQPLRVNVLDTSWTPPPPALREASELLRELCKMMWKYCISIPBKQOLQOXAAY	857
Dd		720	-----EASDDL-----	726
OY		858	ASEIFGKRDYPOSVPRLFISTRLGTEISPRVLOSLS-EPIDYAAPPVKYDRKGYRP	916
Dd		727	-----LNKKERRRSINRNFIQDIYIGME-HPELQQFVGKREKIDFAITYKKYDR-KFG	779
OY		917	RROLLLTPSANVLYEDAKVQO-----RIDYANLTGISVSSLSDSLFVLHVQ	963
Dd		780	VKROLLTLPKCLYLIIGREKVAGQDPKGLVKEVLRKIEIEFILSVLSISTMODDIFILHQ	839
OY		964	REDNKQGDVYLOSDBHVIETLTKT---ALSADRYNNINIINGSTIFAG-----G	1009
Dd		840	-----EYDSLSESYFKTEBLSLKKRYEKTQODPLKFSWTLELTKKENWG	887
OY		1010	PGRDGI---IDFTSG-SELLITTAKNGHAV-VAPRL--NSR	1044
Dd		888	PGVGAGASROYQFHQGFDLAVLAPSKNVLVQVSTGPJPKNSR	930

RESULT 11

QBMMNT7	ID	QBMMNT7	PRELIMINARY:	PRT,	1098 AA.
AC	QBMMNT7:				
DT	01-MAR-2002	(TREMBLrel. 20, Created)			
DT	01-MAR-2002	(TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	MYOsln-1f.				
GN	MYOF.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TextID=9606;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21664126; PubMed=11804589;				
RA	Krugman S., Anderson K.E., Ridley S.H., Rizzo N., McGregor A.,				
RA	Coadwell J., Davidson K., Equinice A., Ellison C.D., Lipp P.,				
RA	Manifava M., Kistakis N., Painter G., Thuring J.W., Cooper M.A.,				

RESULT 13

09NGL7 PRELIMINARY; PRT; 1053 AA.
 ID 09NGL7;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-DEC-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Amoeboid myosin I (Fragment)
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 NC NCB1;taxid=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20440049; PubMed=10984411;
 RA Strookin V., Seipel S., Krendel M., Bonder E.M.;
 RT Characterization of sea urchin unconventional myosins and analysis of
 RT their patterns of expression during early embryogenesis.;
 RL Mol. Reprod. Dev. 57:111-126(2000).
 DR EMBL; AF248488; AAF71717.1;
 DR HSSP; P08799; 1MND.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam; PF00063; myosin_head; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR PRODOM; PD000355; myosin_head; 1.
 DR SMART; SM00242; MYSC; 1.
 FT SMART_TER 1053 1053
 SQ SEQUENCE 1053 AA; 119723 MW; 0B31468B9F9FD03C CRC64;

Query Match 31.4%; Score 1695; DB 5; Length 1053;
 Best Local Similarity 37.4%; Pred. No. 4.9e-108;
 Matches 400; Conservative 171; Mismatches 291; Indels 208; Gaps 24;

24 RDVGVQDFVLENTSFAFIENLRRENNLYTYGTVLVNRYDQIYSRQME 83
 10 RRTSGEDMVLISKI-QGCAIVYENLKRFMDLLTYTIGVLSINPKMPFTEKEVD 68
 84 RRGVSEYVPPHLEFAVDYVRLKTERDQAVMISGESGAGTEATKRLQFYAPTCP 143
 69 MYGGAIIYENPPHYALTDYMYRMNLYEYNGCVIISGESGAGTVAAKIMAYIAKV-- 126
 144 APERGA---VRDLQSNPVLEAFGAKTLRDNSSRPGKYMDVDFEKGAPVGGHIL 199
 127 ---SGGATVQVHKETIOSNPLEAFGNAKTVRNNSRKGVEIDFTKAGAPDGGKIT 184
 200 SYLLEKSRVQNHGNERNFHFYOLLEGEEETRLRLGRNPOSYLYLVKGCACAVSSI 259
 185 NFLLEKSRVQNHGNERNFHFYOLLEGEEETRLRLGRNPOSYLYLVKGCACAVSSI 259
 260 NDKSDMKYMKALSYIDTEDEVDLISIVASVLIHGINFPADEDSNAQVTTENOLKYL 319
 244 DDVEEYKDTLNAHSVVISSEDDQENVLSVAGLHGINVF-VEKIDYAVIHDEFLDPP 302
 320 TRLGVEGTVLEALTRKIIA---KGEELSPILNEOAAYRADLAAYVIRTWTLY 375
 303 SYLGLDKEALRKILRDNSSRPGKYMDVDFEKGAPVGGHIL 319
 376 RKIRSLASKDAESPMSRSTTVGLLDIYGFVEFHNSFEQFNCINCKEQLQDLFETLY 435
 363 SSINGAOKQKVE-----ITIGVLDIYGFVEFHNSFEQFNCINCKEQLQDLFETLY 415
 436 KSEOEYEAGIAMEPEYENNTKIIDLVBEKFKIISLDECL---RPGEAIDLTEL 491
 416 KAEDEEYVAGIEMKATIDYFNKTKIDLVESKVPVPMCLDVCATMAVSGADEKLL 475
 492 EKLEDTVPRPHFLTLADQTKRSLDRGEFRLHYAGEVTVSVGFCDKNDLFRLL 551
 476 OKMNSAVGTHQHY-----GVLSGFLVHHYAAQVRYTVGFEKRRDVLFDL 523
 552 KETWCSSMNPIMACQF-DKSELSDKRPRETVATQFMSSLLQVLEILSRKAPVIRICPN 610

Db 524 IELWSSSENFMTLFPPEKIAGAKSRPTTASSKIKTQANKLVLDLMQCTPHYIRICPN 583
 Qy 611 DAKQGFDEVLNRHOKYILGLEMENRVRAGFAIRKRYEAFLOAKYKSLCPETWMAGR 670
 Db 584 ETRPQDEKRYRHYHVEYILGLENIRVRAGFAIRKRYEAFLOAKYKSLCPETWMAGR 643
 Qy 671 PDGVAVLVNRHLYGKPEEYKMGRTKIFRPTLTATEDSLVRHQSALATKIQAAMRGFH 730
 Db 644 VKSGVNHLLRAVMMQDEROLGKTAVETIKRPSLLEBQRKRFPHARIQKAFQFN 703
 Qy 731 WRKFKRVRSAICIOSMWRGTGLRRKAARRMAAQTIRLLRGLFHLHSPRCENAFEL 790
 Db 704 AOKYVLRKQO-----
 Qy 791 DHVRASFLNLRQLPRNVLDTSWPTPPALREASELRELCKMNMVTKYCRSISPEWKO 850
 Db 715 -----
 Qy 851 QLOQKAVASEIFKKNKNDYQSPRLISTRLGTE-EISPVYLSIGS--EPIDYAVPV 907
 Db 715 -----ASDILFDKERRRYSILNRNFYGDYIGFEHNVS---LQALVARRRIEFADKN 764
 Qy 908 KYDRKGKRPRLQLLTPSAVIY--EDAK-----VKORIDYANLTGISVSSLS 954
 Db 765 KYDRK-FKTKKDLISGRNLYLIGREQIKKQPEKQIRDYVKKRIPLNHISVSLTLQ 823
 Qy 955 DSLFVLHVQREDNKQGVNVLQSDHVIETLT---KTALSADRVNNINING----- 1002
 Db 824 DMFVLHVSEE-----YDLSLESTPEKTEFLVLVKKRPFETIQOLAMNFNDSIFSVKKE 878
 Qy 1003 -----SIFPAGRGNDGIIDFTSGSELLITAKNGHLAVAPRINSR 1044
 Db 879 GNGGCGTTRVRAQSGDPTAIL--KPSKILTVSISG-----LPR-NSR 920

RESULT 14
 061080 PRELIMINARY; PRT; 1186 AA.
 ID 061080;
 AC 061080;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Myosin IC heavy chain.
 DE MICHC.
 OS Acanthamoeba castellanii (Amoeba).
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.
 NC NCB1;taxid=5753;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88016163; PubMed=3477803;
 RA Jung G., Korn E.D., Hammer J.A. III;
 RT "The heavy chain of Acanthamoeba myosin IB is a fusion of myosin-1-like
 RT and non-myosin-like sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:6720-6724(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99079990;
 RA Wang Z.Y., Wang F., Sellers J.R., Korn E.D., Hammer J.A. III;
 RT "Analysis of the regulatory phosphorylation site in Acanthamoeba
 RT myosin IC by using site-directed mutagenesis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:15200-15205(1998).
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL; AF051353; AAC98089.1;
 DR HSSP; P08799; 1MND.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam; PF00063; myosin_head; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR PRODOM; PD000355; myosin_head; 1.
 DR SMART; SM00242; MYSC; 1.

DR SMART: SM00326; SH3: 1.
 DR PROSITE; PS50002; SH3: 1.
 DR SH3 domain.
 SQ SEQUENCE 1186 AA; 129459 MW; E37AD44A685803A6 CRC64;

Query Match 29.1%; Score 1573.5; DB 5; Length 1186;

Best Local Similarity 35.6%; Pred. No. 1.4e-99;
 Matches 373; Conservative 165; Mismatches 304; Indels 207; Gaps 23;

QY 28 GVQDVLLENFTSEAFIEMLRRRFRRENLIYTYIGVLYSVNPRDL-QIYSQHMERFR 86
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 8 GVDDMVLISND-AINDMLKRRFADLIYTYIGHVLSVNPQKQINNLITETKTDYR 66
 QY 87 GVSFEYEPHLEFAVDYVYRALPTEERDOAVMISGESGAGTEATKRLQFYAETCPAPE 146
 | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 DB 67 GKRYELPPHYALADMYRMLSESDQCYIISGESGAGTEATKRLQFYAETCPAPE 126
 QY 147 RGAVERDLQSNPYLEAFNAKTLRNDNSSREKTYNDVOFDFGAVGGHILSYLLEKS 206
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 127 DVMRKVDILSNPLLEAFNAKTLRNNSSRFEGKMEIOFDLKGDPVGGISVYLLEKS 186
 QY 207 RVVHONHGERNFHFYQTLLEGGEDETLRLGLEENPOSYLYVKGQAKYSSINDKSDMK 266
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 187 RVVQTNGERNFHFYQTLLEGGEDETLRLGLEENPOSYLYVKGQAKYSSINDKSDMK 245
 QY 267 VMKRALSYDTEDEVEDLSTIVASVHLGNHFADEDSNAOYTENQKLTLLGVE 326
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 246 DTVMAMKVIYGTAEQHEIFRLYVAILYLVNVOVDGKGSTIADKQVEMLAFLMTE 305
 QY 327 GTTLREALTHKRIAKGEE-----LSPPLNLOAAVAPRALKAVSRFTVLVRRIN 379
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 306 PVAAQALLYR-TITTGQGRGSSVSCPDPLGATSKRALSLRNFYTIQRYN 364
 QY 365 DAMYIDDEPA-----LTGILDFEGEIFGKNGEQCLINFEKLIQIQLTKAEQ 418
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 440 EYEAEGJAMEPVQYFNKKIICDLVEEF-KGIISILDECLRPQEAIDLFLKLEDTV 498
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 419 EYEAEGJOWENIDIVFNKKIICDLLEKRPGLMILTVDCNFP-KGTDDEKREKLGA 477
 QY 499 KHPHFLTKLADQKTRKSLDGRGEPRLHVAEYVSYTGFLDKNNDLIRNLKETMCS 558
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 478 PTHALATSQPD-----EVIHNVAGDYYVNVGFDKKNKDLFLDLIGLACT 527
 QY 559 MNPIMAGCFD--KSELSDKKRPEYVATQFKSLDOLVELLSKEPATIRCIKPNDAKPG 616
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 528 STFFAGLFPPEAKVATSKKPTTAGFKIKESINITVATLSKCTPHYIRCIKPNDAKPG 587
 QY 617 REDEVLIHQVYLYGLMELRVRRAGFAFRKYEAFLOKYSLSCEPTPMNAGRPODGA 676
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 588 DPNNSLVHGVYTLLEENVRIRAGYARQYDDFEFRYRVYVCKTMSGNGMDVSGAE 647
 QY 677 VLVRILGYK-PEEYKMGRTKIFIRPKTILFATEDSLEVRQSLATKIQAMRGFWROKF 735
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 648 AILNHGMSLGEYQGGKTIKIFIRPESVFSLEEDRTVFSYANKIRFLKRTAMRY 707
 QY 736 LAYKRAICIQSMWGTGRRRAKRAKMAQTIRLIRGFIILRHSPROENAFDLHVA 795
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 708 YEYKKG-----NDALVKKERR----- 726
 QY 796 SFLLNLRQDLPRNVLDTSWTPPPALREASELLREICMKNMVKYCRSISPWKQOLOK 855
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 727 ---LSLEP----- 732
 QY 856 AVASELFGKKDNYPQSVRLPISLGLTEISPRVLOSGLSEPIQYAVPYKDK--G 913
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 733 ---FKTDYINRYNFK--LKDCIGK-----GTEKVFALCNLDNLSFNG 773
 QY 914 YKPRPOLLTPSAVIVE-----DAKV-----KQIDYANLIGISVSLDLSL 958
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 774 SKVERRIKVLTSNMLVALIDPKKDIKKKVPFLYVLRKRRIIDFNKIGSITLSPQDNM 833

QY 959 VLHVQREDNKQGDVYLOSDAVIEFTLTKAL-----SADRVNININGSI 1004
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 834 LISTVNGEHS-----NLLECRKRELLGVLLKHPVRIQFADTFN-----V 874
 QY 1005 TRAGG-----PGRDGIIDFTSGS 1022
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 875 TLKGKTCVYKFRDPQSGDKVKGTKYS 903

RESULT 15

ID Q19901 PRELIMINARY; PRT: 1100 AA.

AC Q19901; 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)

DE F29D10.4 protein.

GN F29D10.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

FM [1]

RA SEQUENCE FROM N.A.

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

RP [2]

RA none;

RT none;

RL none;

CC "Genome sequence of the nematode C. elegans: A platform for

investigating biology."

RL Science 282:2012-2018(1998).

CC "SIMILARITY: CONTAINS 1 SH3 DOMAIN.

EMBL: Z75952; CAB00095.1; --

DR HSSP; P08799; LMD.

DR InterPro; IPR000048; IQ_region.

DR InterPro; IPR001609; myosin_head.

DR InterPro; IPR001452; SH3.

DR InterPro; IPR000169; SHprot_acsite.

DR Pfam; PF00612; IQ_1.

DR Pfam; PF00063; myosin_head_1.

DR Pfam; PF00018; SH3_1.

DR PRINTS; PR00193; MYOSINHEAVY.

DR ProDom; PD000066; SH3_1.

DR ProDom; PD000355; myosin_head_1.

DR SMART; SM00326; SH3: 1.

DR PROSITE; PS50002; SH3: 1.

DR PROSITE; PS00639; THIOLEPROTEASE_HIS; UNKNOWN_1.

DR SH3 domain.

SQ SEQUENCE 1100 AA; 124885 MW; 2CDB6F02EDEDED9 CRC64;

Query Match 29.0%; Score 1564; DB 5; Length 1100;

Best Local Similarity 34.7%; Pred. No. 5.7e-99;

Matches 367; Conservative 172; Mismatches 320; Indels 200; Gaps 21;

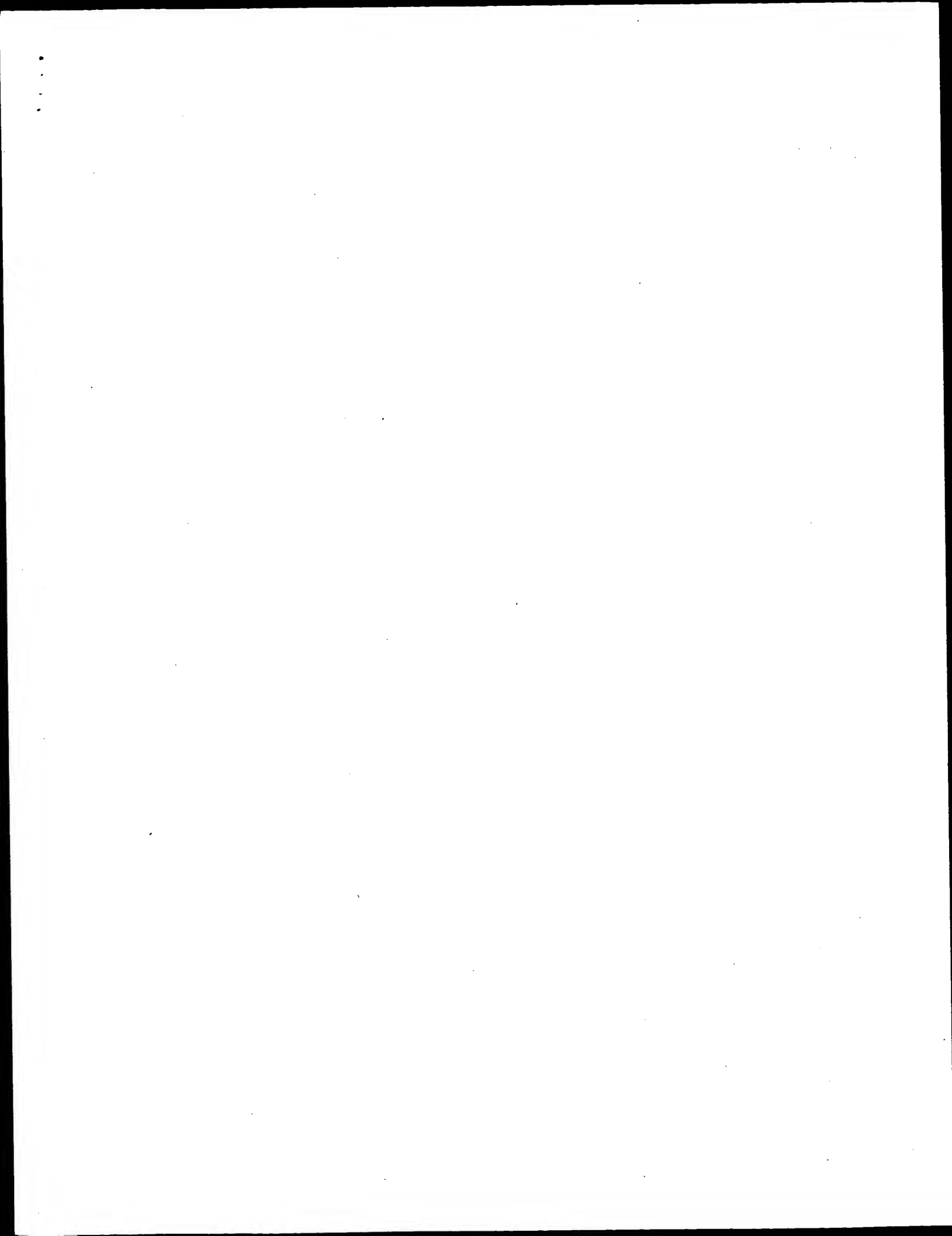
QY 27 GVQDVLLENFTSEAFIEMLRRRFRRENLIYTYIGVLYSVNPRDLQIYSQHMERFR 86
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 14 GVDDMVLISND-AINDMLKRRFADLIYTYIGHVLSVNPQKQINNLITETKTDYR 72
 QY 87 GVSFEYEPHLEFAVDYVYRALPTEERDOAVMISGESGAGTEATKRLQFYAETCPAPE 146
 | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
 DB 73 GAOYENAPHYALADMYRMLSESDQCYIISGESGAGTEATKRLQFYAETCPAPE 132
 QY 147 RGAVERDLQSNPYLEAFNAKTLRNDNSSREKTYNDVOFDFGAVGGHILSYLLEKS 206
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 133 KVQHIKVDILSNPLLEAFNAKTLRNNSSRFEGKMEIOFDLKGDPVGGISVYLLEKS 192
 QY 207 RVVHONHGERNFHFYQTLLEGGEDETLRLGLEENPOSYLYVKGQAKYSSINDKSDMK 266
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 193 RVVHONHGERNFHFYQTLLEGGEDETLRLGLEENPOSYLYVKGQAKYSSINDKSDMK 251

```

QY 267 VMKALSVIDFTEDEVEDLISVAVLHLCNHFHADEDSNAQVTTENOLKYLTRLGVE 326
Db 252 STLHAKVGVVGNODQLEVLRIVAVLHIGNTF-TEENNFPAVSGDKYLEPAPLGLT 310
QY 327 GTTLREALTHRKIIAK-----GELLSPNLNQAAYARDALAKAVSRFTWLVKRNLSL 382
Db 311 SADIEAKLTGKKMESKMGTOKEIDMKLWNEQASYTRDAMVKAIVARLFDYLVKKVNDAM 370
QY 383 ASKDAESPMSRSTTVLGLDIYGFVEFOHNSFQFCINYCNEKLOLFTELTKSEOEY 442
Db 371 ---NITSQSTSDNPSVGIIDYGFELFNNGFEQFCINYNKLOQIFTELTKAQOEY 427
QY 443 EAGCIAMEPQYQNNKIIICDLVEKF-KGIISILDECLR---PGEATDLFTELKEDTV 498
Db 428 VREGIKWTEIDYEDNKIVCLDLETKRPPGIMSLDDTCAQNHGQREGVDRQLTLTSKSF 487
QY 499 KPHHFLTHKLADQKTRKSLDRGEFRLHAGVUTYSVGFGLDKNNDLFRNLKETWCS 558
Db 488 AGHPHF-----GPGSDSFYKHYAGDVYINVDGFCDRNRDVLYPDLLIMOKS 535
QY 559 MNPIMACCEDEKS-ELSDKRPETVANOFKMSLQLEILRSKEPAYIRCIRPNDAKOPR 617
Db 536 SREPIQALFEPENVAASAGKRPJTSTKIRTOANTLVESELMKSPHYVRCIKNETKRPND 595
QY 618 FDEVLIRHQQYKYLGLMENLRVRAAGAVRKTEAELOKYSCLCPETWPMAGRPQGVAV 677
Db 596 WEBSRVKHOVEYGLKRENI RVRAGAFYRABDKFAORAYIVSPQWPCFQDQORACEI 655
QY 678 LVHNLGKPEEYKMGRTKIFIRFPKTLFATEDSLVBRQSLATKIOAMRGPHMRQKFLR 737
Db 656 ICDSVHMEKNQYOMGKTKIFVKNPESLFEETREKFDGAYRIQKAMROFSAR----- 710
QY 738 VKRSAICIOSMWRGTLGRKKAARKWAQTIIRLLRGFTLRHSPRCPENAFFLDHVRASF 797
Db 711 ----- 710
QY 798 LNLRLRQLPRNVLDTSWTPPPALREASBLRELCKMNMWYKCSISPEWKOLOQKAV 857
Db 711 -----KOHIKOEQ 719
QY 858 ASEIEFKGKNDYPOSVPLFISTRLGTEESPRVLOSL--GSEPIQAVPVVKYDRKGYK 915
Db 720 AADLMYGKKERRRYSLNRFVGDYIGLEH--HPTLOSLVGKROYLPACTANKYDKR-FR 776
QY 916 PRPRLTLTPSAVIVEDAKV-----KQRIYANLTGISVSLSDSLFVLHV 962
Db 777 VKRLDLLLVNHLTLIGKEKYNKNGPEKGIYEVIRQFDLPQIKSIGISPYQDDFYLYL 836
QY 963 OREDNKQKGDVYLOSDHIEFTLK---TALS---ADRYNN-----ININGST 1004
Db 837 GNDYSS-----LETPFKTEFCTALSKAVKERTNGTLHLDRSSHVSYYKKMKF 886
QY 1005 TTAGGP-----GRDGIIDFTSGSELLITKAKNGHLAVVA 1038
Db 887 DFSDGKRTYQFCGNDG---TSSAEK--TLKPNKGVNVS 919

```

Search completed: July 7, 2003, 14:23:55
 Job time : 114.37 secs



Run on: July 7, 2003, 14:11:39 ; Search time 1.53962 Seconds

1384.761 Million cell updates/sec

Title: US-09-893-371-2

Sequence: 1 MRYRASALGSDGVRVT 16

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

1: _Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	62.8	262	20	AAV25521	E. coli mutant FabH
2	49	62.8	262	22	AAU34559	E. coli cellular F
3	49	62.8	262	22	AAE02202	Escherichia coli N
4	49	62.8	269	22	AAU38231	Salmonella typhi c
5	49	62.8	429	20	AAV25520	E. coli FabI protease
6	49	62.8	491	22	ABG28562	Novel human diacyl
7	49	62.8	693	22	ABG28564	Novel human diacyl
8	48	61.5	267	22	AAB90827	Human shear stress
9	48	61.5	537	22	AAAM4773	Human polypeptide
10	48	61.5	1050	22	ABG10171	Novel human diacyl

45	11	48	61.5	1.063	22	AAG39929
44	12	45	57.7	3.57	21	AAK16522
43	13	44	56.4	2.65	22	AAU35255
42	14	44	56.4	2.68	18	AAW98083
41	15	41	52.6	4.15	23	ABB93285
40	16	40	51.3	5.26	10	AAAP90587
39	17	40	51.3	3.32	22	AAO11666
38	18	40	51.3	4.6	22	ABP05139
37	19	40	51.3	5.1	23	ABP05815
36	20	40	51.3	4.02	23	AAU75522
35	21	40	51.3	4.02	23	AAU75522
34	22	40	51.3	4.25	22	ABBS24983
33	23	39	50.0	6.4	22	AAO08242
32	24	39	50.0	7.0	22	AAU27464
31	25	39	50.0	18.3	21	AAO08315
30	26	39	50.0	2.24	20	AAAG58516
29	27	39	50.0	2.61	21	AAW97727
28	28	39	50.0	2.61	21	AAAB03652
27	29	39	50.0	2.61	21	AAAB03653
26	30	39	50.0	2.62	21	AAAB03636
25	31	39	50.0	2.62	23	ABBA8921
24	32	39	50.0	2.79	21	AAAG40867
23	33	39	50.0	3.15	21	AAAG40866
22	34	39	50.0	3.25	21	AAAG40865
21	35	39	50.0	2.92	21	AAAB07665
20	36	38	48.7	4.1	22	AAO13360
19	37	38	48.7	4.6	22	AAO09624
18	38	38	48.7	1.23	21	AAAG47869
17	39	38	48.7	1.44	21	AAAG47868
16	40	38	48.7	1.57	19	AAAG62730
15	41	38	48.7	1.70	21	AAAG47867
14	42	38	48.7	2.18	21	AAAG43357
13	43	38	48.7	2.19	21	AAAG47863
12	44	38	48.7	2.39	21	AAAG43366
11	45	38	48.7	2.39	21	AAAG43366

ALIGNMENTS

XX	AA25521	standard; Protein; 262 AA.
AC	AA25521;	
XX	30-SEP-1999	(first entry)
DT	E. coli mutant FabI protein.	
DE		
KW	FabI; enoyl ACP reductase; acyl carrier protein; antimicrobial; therapy;	
KW	fatty acid biosynthesis; detection; antibacterial; antifungal; treatment;	
KW	antiprotzoal; antiparasitic; antiviral; soap; infection; disinfectant;	
KW	detergent; mouthwash; toothpaste; contraceptive; inhibitor; triclosoan;	
KX	resistance; NSAM; non-specific antimicrobial; mutant.	
OS	Escherichia coli.	
XX	Synthetic.	
Key	Location/Qualifiers	
FT	1..262	
FT	/label= FabI	
FT	/note= "No stop codon given in specification"	
PN	WO9937800-A1.	
PD	23-JUL-1999.	
PF	22-JAN-1999;	99WO-US01288.
PR	26-JAN-1998;	98US-0013440.
PR	23-JAN-1998;	98US-0072244.

PA (TUFT) TUFTS COLLEGE.
 XX
 PI Levy SB, McNurry LM;
 XX
 DR WPI: 1999-458702/38.
 DR N-PSDB: AAX88348.
 XX
 PT Identifying antimicrobials that target enoyl acyl carrier protein
 PT reductase, for therapeutic use and for incorporation into e.g. soaps
 PT and detergents
 XX
 PS Claim 77; Fig 2; 80pp; English.
 XX
 CC This invention describes a novel method for identifying antimicrobials
 CC (1) that interact with an enoyl acyl carrier protein (ACP) reductase
 CC (ER) polypeptide, which is essential for fatty acid biosynthesis, by
 CC contacting ER with test compound and detecting an interaction. The
 CC identified antimicrobials have antibacterial, antifungal, antiparasitic,
 CC antiparasitic and antiviral activity. The products of the invention are
 CC used (1) therapeutically to treat a wide variety of viral, bacterial,
 CC fungal, yeast and protozoal infections, in both humans and animals, and
 CC (11) in disinfectants, detergents, soaps, mouthwashes, toothpaste's and
 CC also contraceptive devices. The invention describes methods for (1)
 CC identifying a (1) from its effect on enzymatic activity of ER, (2)
 CC identifying a (1) from its ability to inhibit fatty acid biosynthesis in
 CC a microbe, (3) identifying a (1) that interacts with a mutant ER, (4)
 CC identifying (1) that inhibit proliferation or viability of a microbe that
 CC is resistant to triclosan or to NSAM (non-specific antimicrobials). This
 CC sequence represents a mutant Escherichia coli FabI protein which
 CC interacts with the enoyl-acyl carrier protein (ACP) reductase to
 CC illustrate the method of the invention.
 XX
 SQ Sequence 262 AA:
 XX
 Query Match 62.8%; Score 49; DB 20; Length 262;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 DB 170 VRYMANAMGPEGVRV 184
 RESULT 2
 AAU34559
 ID AAU34559 standard; Protein: 262 AA.
 XX
 AC AAU34559;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE E. coli cellular proliferation protein #140.
 XX
 KM Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.
 XX
 OS Escherichia coli.
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.

XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR N-PSDB: AAS52418.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprising sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 10152; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 262 AA:
 XX
 Query Match 62.8%; Score 49; DB 22; Length 262;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 DB 170 VRYMANAMGPEGVRV 184
 RESULT 3
 AAE02203
 ID AAE02203 standard; Protein: 262 AA.
 XX
 AC AAE02203;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Escherichia coli NADPH-dependent enoyl-ACP reductase (FabI).
 XX
 KM FabI: high throughput method; fatty acid biosynthesis; therapy;
 KM bacterial enzyme; biological agent screening; otitis media; empyema;
 KM bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;
 KM infective endocarditis; secretory diarrhoea; splenic; retroperitoneal;
 KM intrarenal; perinephric; cerebral; cutaneous; abscess; blepharitis;
 KM conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;
 KM cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndrome;
 KM impetigo; folliculitis; wound infection; bacterial myositis;
 KM septic arthritis; osteomyelitis; enoyl-ACP reductase;
 KM acyl carrier protein.
 XX
 OS Escherichia coli.
 PN WO200130968-A1.
 XX
 PD 03-MAY-2001.
 XX
 PR 26-OCT-2000; 2000WO-US29451.
 PR 27-OCT-1999; 99US-0161775.
 XX
 PA

XX (SMIR) SMITHKLINE BEECHAM CORP.
 PA (SMIR) SMITHKLINE BEECHAM PLC.
 XX
 XX
 PI Dewolf W, Kallender H, Lonsdale JT;
 DR WPI; 2001-316332/33.
 DR N-PSDB; AAD06214.
 XX
 PT High throughput method for screening for biological agents against
 PT fatty acid biosynthesis comprises contacting a bacterial enzymatic
 PT pathway with enzymes e.g. malonyl-CoA:ACP transacylase -
 PS
 PS Claim 1; Page 15; 94pp; English.
 XX
 CC The present invention relates to a high throughput method for screening
 CC biological agents affecting fatty acid biosynthesis, comprises
 CC contacting a bacterial enzymatic pathway with enzymes. The method is
 CC used for screening biological agents affecting fatty acid biosynthesis.
 CC Agonists and antagonists of Fab (fatty acid biosynthesis) are used to
 CC inhibit, prevent or treat diseases such as infections of the upper
 CC respiratory tract (e.g. otitis media, bacterial tracheitis, acute
 CC epiglottitis, thyroiditis), lower respiratory (e.g. empyema, lung
 CC abscess), cardiac (e.g. infective endocarditis), gastrointestinal (e.g.
 CC secretory diarrhoea, splenic abscess, retroperitoneal abscess), central
 CC nervous system (e.g. cerebral abscess), eye (e.g. blepharitis,
 CC conjunctivitis, keratitis, endophthalmitis, preseptal and orbital
 CC cellulitis, dacryocystitis), kidney and urinary tract (e.g.
 CC epididymitis, intrarenal and perinephric abscess, toxic shock syndrome),
 CC skin (e.g. impetigo, folliculitis, cutaneous abscess, cellulitis, wound
 CC infection, bacterial myositis) and bone and joint (e.g. septic
 CC arthritis, osteomyelitis).
 CC The present sequence is Escherichia coli NADPH-dependent
 CC enoyl-ACP (acyl carrier protein) reductase (FabI). In fatty acid
 CC biosynthetic pathway, dehydration by FabZ leads to trans-2-enoyl-ACP
 CC which is in turn converted to acyl-ACP by FabI.
 CC
 SO Sequence 262 AA:
 Query Match 62.8%; Score 49; DB 22; Length 262;
 Best Local Similarity 60.0%; Pred. No. 2.2;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 Db 170 VRYMANMGPEGVRV 184
 RESULT 4
 AAU38231
 ID AAU38231 standard; Protein: 269 AA.
 AC AAU38231;
 XX
 DT 14-FEB-2002 (first entry)
 DE
 XX
 DE Salmonella typhi cellular proliferation protein #122.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 OS Salmonella typhi.
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS56090.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 PS
 PS Example 3; Seq ID No 13824; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 269 AA:
 Query Match 62.8%; Score 49; DB 22; Length 269;
 Best Local Similarity 60.0%; Pred. No. 2.3;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRYRASALGSDGVRV 15
 :|||:|:|:|:|:|
 Db 177 VRYMANMGPEGVRV 191
 RESULT 5
 AAY25520
 ID AAY25520 standard; Protein: 429 AA.
 AC AAY25520;
 XX
 DT 30-SEP-1999 (first entry)
 DE
 XX
 DE E. coli FabI protein fragment.
 XX
 KW FabI; enoyl ACP reductase; acyl carrier protein; antimicrobial; therapy;
 KW fatty acid biosynthesis; detection; antibacterial; antifungal; treatment;
 KW antiprotozoal; antiparasitic; antiviral; soap; infection; disinfectant;
 KW detergent; mouthwash; toothpaste; contraceptive; inhibitor; triclosan;
 KW resistance; NSAM; non-specific antimicrobial.
 OS Escherichia coli.
 PN
 XX
 FH Key Location/Qualifiers
 FT Protein 1..429
 FT /label= FabI
 FT /note= "No start or stop codon given in specification"
 FT Misc-difference 74
 FT /note= "in-frame stop codon encoded by TGA"
 FT Misc-difference 111

	P	/note= "In-frame stop codon encoded by TAA"
FT	Misc-difference	393 /note= "In-frame stop codon encoded by TAA"
FT	Misc-difference	398 /note= "in-frame stop codon encoded by TAA"
XX		
PN	WO9937800-A1.	
XX		
PD	29-JUL-1999.	
XX		
PF	22-JAN-1999;	99MO-USO1288.
XX		
PR	26-JAN-1998;	98US-0013440.
PR	23-JUN-1998;	98US-0072244.
XX	(TUFT) TUFTS COLLEGE.	
PA		
XX		
P1	Levy SB, McMurry LM;	
DR	NPI; 1999-458702/38.	
DR	N-PADB; AAX88347.	
PT	Identifying antimicrobials that target enoyl acyl carrier protein reductase, for therapeutic use and for incorporation into e.g. soaps and detergents	
PS		
PS	Example 5; Page 68-70; 80pp; English.	
CC	This invention describes a novel method for identifying antimicrobials (I) that interact with an enoyl acyl carrier protein (ACP) reductase (ER) polypeptide, which is essential for fatty acid biosynthesis, by contacting ER with test compound and detecting any interaction. The identified antimicrobials have antibacterial, antifungal, antiprotozoal, anti-parasitic and antiviral activity. The products of the invention are used (i) therapeutically to treat a wide variety of viral, bacterial, fungal, yeast and protozoan infections, in both humans and animals, and (ii) in disinfectants, detergents, soaps, mouthwashes, toothpaste's and also contraceptive devices. The invention describes methods for (1)	
CC	identifying a (I) from its effect on enzymatic activity of ER, (2)	
CC	identifying a (I) from its ability to inhibit fatty acid biosynthesis in a microbe, (3) identifying a (I) that interacts with a mutant ER, (4)	
CC	identifying (I) that inhibit proliferation or viability of a microbe that is resistant to triclosen or to NSAM (non-specific antimicrobials). This sequence represents the Escherichia coli FabI protein which interacts with the enoyl-acyl carrier protein (ACF) reductase to illustrate the method of the invention.	
SQ	Sequence 429 AA:	
	Query Match	Score 49; DB 20; Length 429;
	Best Local Similarity	60.0%; Pred. No. 3.7;
	Matches 9; Conservative	4; Mismatches 2; Indels 0; Gaps 0;
OY	1 MKRYASALGSDGV RV 15 :: : : DB 300 VRYMANMAGPEGRV 314	
ID	ABG28562 standard; Protein; 491 AA.	
XX		
AC	ABG28562;	
DT	18-FEB-2002 (first entry)	
DE	Novel human diagnostic protein #28553.	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder. Homo sapiens.	
OS	Homo sapiens.	
XX		

XX	PN	WO200175067-A2.
XX	PD	11-OCT-2001.
XX	PX	30-MAR-2001; 2001MO-US08631.
PF	PR	31-MAR-2000; 2000US-0540217.
XX	PR	23-AUG-2000; 2000US-0649167.
XX	PA	(HYSE-) HYSEQ INC.
XX	PI	Drmancic RT, Liu C, Tang YT;
XX	DR	WPI: 2001-639362/73.
XX	DR	N-Psdb; AAS92749.
PT	PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	PT	diagnostics, forensics, gene mapping, identification of mutations
PT	PT	responsible for genetic disorders or other traits and to assess
PT	PT	biodiversity -
PS	Claim 20;	SEQ ID NO 58921; 103pp; English.
XX	CC	The invention relates to isolated polynucleotide (I) and
CC	CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	CC	and gene mapping, and in recombinant production of (II). The
CC	CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	CC	to restore normal activity of (II) or to treat disease states involving
CC	CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	CC	a food supplement. (II) and its binding partners are useful in medical
CC	CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	CC	disorders involving aberrant protein expression or biological activity.
CC	CC	The polypeptide and polynucleotide sequences have applications in
CC	CC	diagnostics, forensics, gene mapping, identification of mutations
CC	CC	responsible for genetic disorders or other traits to assess biodiversity
CC	CC	and to produce other types of data and products dependent on DNA and
CC	CC	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	CC	diagnostic amino acid sequences of the invention.
CC	CC	Note: The sequence data for this patent did not appear in the printed
CC	CC	specification, but was obtained in electronic format directly from WIPO
CC	CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	Sequence	491 AA;
SQ	Query Match	62.8%; Score 49; DB 22; Length 491;
	Best Local Similarity	60.0%; Pred. No. 4.3;
	Matches 9; Conservative	4; Mismatches 2; Indels 0; Gaps 0.
OY	1 KRRYSALGSDGVRY 15	: :: :
DB	378 VRYMANMGPESVRY 392	:
RESULT 7		
ABG28564		
ID	ABG28564 standard; Protein: 693 AA.	
AC	ABG28564;	
DT	18-FEB-2002 (first entry)	
DE	Novel human diagnostic protein #28555.	
XX		
KW	Human: Chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement; medical imaging; diagnostic; genetic disorder.	
OS	Homo sapiens.	
XX		
PN	WO200175067-A2.	
XX		

PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI: 2001-639362/73.
DR N-PSDB: AAS92751.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20: SEQ ID No 58923; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.
XX
XX SQ Sequence 693 AA;
XX
XX Query Match 62.8%; Score 49; DB 22; Length 693;
XX Best Local Similarity 60.0%; Pred. No. 6.2;
XX Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 MRYRASALGSDGVRV 15
XX :|||:|:|:|:|:|
XX 601 VRYMANMNGPEGVRV 615
XX
XX RESULT 8
XX AAB90827
XX ID AAB90827 standard; Protein: 267 AA.
XX
XX AC AAB90827;
XX
XX 15-JUN-2001 (first entry)
XX
XX Human shear stress-response protein SEQ ID NO: 171.
XX
XX Human; shear stress-response protein; vascular disease;
XX arteriosclerosis.
XX
XX Homo sapiens.
XX OS
XX WO200125427-A1.
XX PN
XX 12-APR-2001.
XX PD

PF 02-OCT-2000; 2000WO-JP06840.
XX
XX 01-OCT-1999; 99JP-0280976.
XX
XX (KYOW) KYOWA HAKKO KOSYO KK.
XX (NOJ1) NOJIMA H.
XX
XX Nojima H, Yoshitue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI Kuga T, Sekine S, Nakamura Y, Sugano S;
XX WPI: 2001-266308/27.
DR N-PSDB: AAH02967.
XX
XX DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis -
XX
XX Claim 35: Page 664-665; 678pp; Japanese.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension.
XX
XX SQ Sequence 267 AA;
XX
XX Query Match 61.5%; Score 48; DB 22; Length 267;
XX Best Local Similarity 100.0%; Pred. No. 3.3;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 7 ALGSDGVRV 16
XX :|||||
XX 26 ALGSDGVRV 35
XX
XX DB
XX
XX RESULT 9
XX AAM41777
XX ID AAM41777 standard; Protein: 537 AA.
XX
XX AC AAM41777;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 6708.
XX
XX DE
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX OS
XX Homo sapiens.
XX PN WO200153312-A1.
XX
XX PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX PF 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AAI60933.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 6708; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with neotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemoradio/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 537 AA;
Query Match 61.5%; Score 48; DB 22; Length 537;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ALGSDGVRVT 16
DB 60 ALGSDGVRVT 69

RESULT 10
ID AAG10171 standard; Protein; 1050 AA.
XX
AC AAG10171;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #10162.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT.
XX
DR WPI, 2001-639362/73.
DR N-PSDB; AAS74358.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
XX
PS Claim 20; SEQ ID NO 40530; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAG00010-AAG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1050 AA;
Query Match 61.5%; Score 48; DB 22; Length 1050;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ALGSDGVRVT 16
DB 12 ALGSDGVRVT 21

RESULT 11
ID AAM39991 standard; Protein; 1063 AA.
XX
AC AAM39991;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3136.
XX
KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0635450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX

Journal of Management Inquiry 22(4) 401-417

PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145227.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145911.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 25-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 57.7%; Score 45; DB 21; Length 357;
 Best Local Similarity 69.2%; Pred. No. 15;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 RASALGSDGYRVT 16
 DB 43 RASALGNTGLKVT 55

RESULT 13
 AAU36258
 ID AAU36258 standard; Protein: 265 AA.
 AC AAU36258;
 XX 14-FEB-2002 (first entry)
 DT XX
 DE Pseudomonas aeruginosa cellular proliferation protein #248.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207227P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR N-PDB: AAS54117.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX

PS Example 3; Seq ID NO 11851; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 265 AA;

SO

Query Match 56.4%; Score 44; DB 22; Length 265;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 173 VRYLAGSLGAGTRV 187

RESULT 14
AAW99839
ID AAW99839 standard; Protein; 268 AA.

XX AAW99839;

AC AAW99839;

XX

DT 08-JUN-1999 (first entry)

XX

DE L. pneumophila Fabi enoyl reductase homologue protein sequence.

XX

KW *Legionella pneumophila*; Fabi; enoyl reductase; antimicrobial;
KW *gemfibrozil*; growth inhibition; bacterium; infection.

XX

OS *Legionella pneumophila*.

XX

PN MO9731530-AI.

XX

PD 04-SEP-1997.

XX

PF 28-FEB-1997; 97WO-US03158.

XX

PR 29-FEB-1996; 96US-0608712.

XX

PA (UYCO) UNIV COLUMBIA NEW YORK.

XX

PI Della-Latta P, Kabbash C, Shuman HA, Silverstein SC;
XX
XX WPI: 1997-448377/41.
XX N-PSDB: AAX19775.

DR

XX

PT Inhibiting growth of bacteria - by contacting the bacteria with
PT *gemfibrozil* or related compound

XX

PS Example 2; Fig 21B; 109pp; English.

XX

XX A method has been developed for: (i) inhibiting growth of a bacterium;
CC (ii) alleviating the symptoms of a bacterial infection in a subject;
CC (iii) inhibiting activity of enoyl reductase enzyme in a cell; or (iv)
CC altering a pathway of fatty acid synthesis in a bacterium, comprises
CC contacting the bacterium, subject or cell with a novel compound (I) or

CC its salt or ester. The above processes are especially applicable to
CC bacteria/bacterial cells selected from *Legionella pneumophila*,
CC *Mycobacterium tuberculosis*, *Bacillus subtilis*, *Bacillus megaterium*,
CC *Pseudomonas oleovorans*, *Alcaligenes eutrophus*, *Rhodococcus* sp.,
CC *Citrobacter freundii*, Group A *Streptococcus* sp., *Coag neg Staphylococcus*
CC *aureus* or *Neocardia* sp. The bacterial infection is e.g. leprosy, brucella
CC or *salmonella*. The present sequence represents a Fabi enoyl reductase
CC homologue from *L. pneumophila*, which is used in an example from the
CC present invention.

XX

SO Sequence 268 AA;

Query Match 56.4%; Score 44; DB 18; Length 268;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 177 VRYLAGSLGSRGLRI 191

RESULT 15
ABB93255
ID ABB93255 standard; Protein; 415 AA.

XX ABB93255;

AC ABB93255;

XX

DT 31-MAY-2002 (first entry)

XX

DE Herbicidally active polypeptide SEQ ID NO 2466.

XX

KW Herbicidal; plant; agriculture; herbicide.

XX

OS Arabidopsis thaliana.

XX

PN WO200210210-A2.

XX

PD 07-FEB-2002.

XX

PF 28-AUG-2001; 2001WO-EP09892.

XX

PR 28-AUG-2001; 2001WO-EP09892.

XX

PA (FARB) BAYER AG.

XX

PI Tietjen K, Weidler M;
XX
XX WPI: 2002-269010/31.

DR

XX

PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -

XX

PS Claim 5; SEQ ID NO 2466; 261pp + Sequence Listing; English.

XX

XX The invention relates to identifying target proteins
CC aligning and comparing nucleic acid or amino acid sequences, comprising
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

XX

SO Sequence 415 AA;

Query Match 52.6%; Score 41; DB 23; Length 415;
Best Local Similarity 53.3%; Pred. No. 82;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 16

Mon Jul 7 15:30:57 2003

us-09-893-371-2.rag

Page 10

Db 392 RYKAGALGAERKRAT 406

Search completed: July 7, 2003, 14:21:22
Job time : 3.53962 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:19:35 ; Search time 0.543396 Seconds
(without alignments)
866.342 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78
Sequence: 1 MRYRASALGSDGVRRV 16

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	62.8	262	1	US-08-241-766-8
2	49	62.8	262	1	US-08-241-766-9
3	38	48.7	308	4	US-09-347-803-12
4	38	48.7	308	5	PCT-US96-03916-6
5	38	48.7	985	5	PCT-US96-03916-6
6	37	47.4	309	4	US-09-134-001C-3039
7	37	47.4	623	3	US-09-041-991A-6
8	37	47.4	943	4	US-09-397-885-5
9	36	46.2	260	4	US-09-134-001C-4009
10	36	46.2	537	1	US-08-173-508-2
11	36	46.2	537	2	US-08-265-310-2
12	36	46.2	537	3	US-08-951-742-2
13	36	46.2	635	3	US-09-041-991A-4
14	36	46.2	702	6	5240838-15
15	36	46.2	3729	2	US-08-804-227C-4
16	35	44.9	48	2	US-08-637-759B-480
17	35	44.9	48	3	US-08-871-355A-480
18	35	44.9	48	4	US-09-201-945-480
19	35	44.9	256	4	US-09-292-412-2
20	35	44.9	303	4	US-09-420-786A-3
21	35	44.9	308	4	US-09-347-803-10
22	35	44.9	338	4	US-09-199-637A-47
23	35	44.9	350	4	US-08-961-083-126
24	35	44.9	5588	4	US-09-036-987A-6
25	35	44.9	5588	4	US-09-370-700-6
26	34	43.6	77	4	US-09-465-558-52
27	34	43.6	113	1	US-08-211-202-114

28	34	43.6	303	2	US-08-286-819A-23	Sequence 23, Appl
29	34	43.6	303	3	US-08-980-357-23	Sequence 23, Appl
30	34	43.6	319	2	US-08-795-927-4	Sequence 2, Appl
31	34	43.6	329	3	US-09-022-669-2	Sequence 4, Appl
32	34	43.6	329	4	US-09-262-749-2	Sequence 2, Appl
33	34	43.6	329	4	US-09-603-567-2	Sequence 2, Appl
34	34	43.6	359	4	US-09-722-139-4	Sequence 4, Appl
35	34	43.6	359	4	US-09-721-832-4	Sequence 4, Appl
36	34	43.6	359	4	US-09-721-832-4	Sequence 4, Appl
37	34	43.6	433	4	US-09-199-637A-349	Sequence 349, App
38	34	43.6	524	1	US-08-529-654-4	Sequence 4, Appl
39	34	43.6	524	3	US-08-957-302A-2	Sequence 2, Appl
40	34	43.6	524	4	US-09-542-403-2	Sequence 2, Appl
41	34	43.6	524	4	US-09-042-709A-20	Sequence 20, Appl
42	34	43.6	623	1	US-08-653-740-7	Sequence 7, Appl
43	34	43.6	623	2	US-09-073-594-7	Sequence 7, Appl
44	34	43.6	623	3	US-09-275-925-7	Sequence 7, Appl
45	34	43.6	633	3	US-09-041-991A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-241-766-8
; Sequence 8, Application US/08241766
; Patent No. 568650

GENERAL INFORMATION:

APPLICANT: JACOBS, W. R.

APPLICANT: COLLINS, D. M.

APPLICANT: BANERJEE, A. M.

APPLICANT: DELISTE, G. W.

APPLICANT: WILSON, T. M.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING

TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN LHA AGENT

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & ROESTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/241,766

FILING DATE: 12-MAY-1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MONROY, GLADYS H.

REGISTRATION NUMBER: 32,430

REFERENCE/DOCKET NUMBER: 25237-20003.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 262 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-241-766-8

Query Match

Best Local Similarity

Matches 9; Conservative

Matches 1 MRYRASALGSDGVRRV 15

62.8%; Score 49; DB 1; Length 262;

Pred. NO. 0.51;

Indels 0; Gaps 0;

Db :|||:|:|:|
170 VRYMANMGPEGVY 184

RESULT 2

US-08-241-766-9
; Sequence 9, Application US/08241766
; Patent No. 5686590
; GENERAL INFORMATION:
; APPLICANT: JACOBS, W. R.
; APPLICANT: COLLINS, D. M.
; APPLICANT: BANERJEE, A.
; APPLICANT: DELISLE, G. W.
; APPLICANT: WILSON, T. M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN INHA AGENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,766
; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADIS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20003.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-241-766-9
Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.51;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 MRYRASNAGSDGVRV 15
Db 170 VRYMANMGPEGVY 184
RESULT 3
US-09-347-803-12
; Sequence 12, Application US/09347803
; Patent No. 6274379
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
; FILE REFERENCE: BB-1176
; CURRENT APPLICATION NUMBER: US/09/347,803
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,952
; EARLIER FILING DATE: July 15, 1998

; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-347-803-12

Query Match 48.7%; Score 38; DB 4; Length 308;
Best Local Similarity 48.9%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ASALGSDGV 13
Db 120 ASALGDDGV 128

RESULT 4

PCT-US96-03916-6
; Sequence 6, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-03916-6
Query Match 48.7%; Score 38; DB 5; Length 985;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 3 YRASALGSDGVRV 16
Db 921 YECTVLISDGRV 934
RESULT 5
PCT-US96-03916-66
; Sequence 66, Application PC/TUS9603916
; GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03916-66

Query Match 48.7%; Score 38; DB 5; Length 985;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 YRSALGSDGVRVT 16
DB 921 YECTVLISDGRVT 934

RESULT 6
US-09-134-001C-3039
Sequence 3039, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3039
LENGTH: 309
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3039

Query Match 47.4%; Score 37; DB 4; Length 309;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRSALGSDGVRV 15
DB 272 RYHTALGTDPVKL 285

RESULT 7
US-09-041-991A-6
Sequence 6, Application US/09041991A
Patent No. 6107278
GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Narva, Kenneth E.
APPLICANT: Muller-con, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,991A
FILING DATE: 13-MAR-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-709
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 623 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-041-991A-6

Query Match 47.4%; Score 37; DB 3; Length 623;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RASALGSDGVRVT 16
DB 539 RVSSIGSSTIRVT 551

RESULT 8
US-09-397-885-5
Sequence 5, Application US/09397885
Patent No. 6323007
GENERAL INFORMATION:
APPLICANT: Moller, Soren
APPLICANT: Johansen, Charlotte
APPLICANT: Schaefer, Thomas
APPLICANT: Ostergaard, Peter Rahbek
APPLICANT: Hoeck, Lisbeth Hedegaard
TITLE OF INVENTION: A 2,6,-D-Fructan Hydrolase Enzyme And
FILE REFERENCE: 5540.200-US
CURRENT APPLICATION NUMBER: US/09/397,885
CURRENT FILING DATE: 1999-09-17
EARLIER APPLICATION NUMBER: PA 1998 01173

EARLIER FILING DATE: 1998-09-18
EARLIER APPLICATION NUMBER: PA 1998 01623
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: 60/101,615
EARLIER FILING DATE: 1998-09-24
EARLIER APPLICATION NUMBER: 60/111,675
EARLIER FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 943
TYPE: PRF
ORGANISM: Paenibacillus macerans
US-09-397-885-5

Query Match 47.4%; Score 37; DB 4; Length 943;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 MYRASALGSDG 12
:1:1: 1:1:
Db 270 LRRRANGOGTGD 281

RESULT 9
US-09-134-001C-4009
Sequence 4009, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GRC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4009
LENGTH: 260
TYPE: PRF
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4009

Query Match 46.2%; Score 36; DB 4; Length 260;
Best Local Similarity 46.7%; Pred. No. 88;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MYRASALGSDGVRV 15
:1:1: 1:1:
Db 175 VKYIALDLGHDNIRV 189

RESULT 10
US-08-173-508-2
Sequence 2, Application US/08173508
Patent No. 5616485
GENERAL INFORMATION:
APPLICANT: Bartfield, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 508
APPLICATION NUMBER: US/08/173,508
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-508-2

Query Match 46.2%; Score 36; DB 1; Length 537;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MYRASALGSDGVRVT 16
:1:1: 1:1:
Db 6 IRRRAFATGALVYT 21

RESULT 11
US-08-265-310-2
Sequence 2, Application US/08265310
Patent No. 5856166
GENERAL INFORMATION:
APPLICANT: Bartfield, Daniel
APPLICANT: Butler, Michael J.
APPLICANT: Hadary, Dany
APPLICANT: Jenish, David
APPLICANT: Krieger, Timothy
APPLICANT: Malek, Lawrence T.
APPLICANT: Soostmeyer, Gisela
APPLICANT: Walczyk, Eva
APPLICANT: Krygsman, Phyllis
APPLICANT: Garven, Shelia
TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 310
APPLICATION NUMBER: US/08/265,310
FILING DATE: 24-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,508
FILING DATE: 23-DEC-1993

US-08-951-742-2

APPLICATION NUMBER: US/
FILING DATE: 24-SEP-1990

FILED DATE: 24-SEP-1
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 300,211
 FILING DATE: 23-JAN-1989
 APPLICATION NUMBER: 759,315
 FILING DATE: 26-JUL-1985
 SEQ ID NO:15
 LENGTH: 702
 5240838-15

Query Match 46.2%; Score 36; DB 6; Length 702;
 Best Local Similarity 77.8%; Pred. No. 2.6e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 SALSGDVR 14
 :||| |||
 Db 277 NALGEDVR 285

RESULT 15
 US-08-804-227C-4
 ; Sequence 4, Application US/08804227C
 ; Patent No. 5876991
 ; GENERAL INFORMATION:
 ; APPLICANT: Dehoff, Bradley S.
 ; APPLICANT: Rostock, Paul R., Jr.
 ; APPLICANT: Sutton, Kimberly L.
 ; TITLE OF INVENTION: POLYPEPTIDE SYNTHASE GENES
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: THOMAS G. PLANT 1501
 ; STREET: LILLY CORPORATE CENTER
 ; CITY: INDIANAPOLIS
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: ASCII(DOS) Text only
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,227C
 ; FILING DATE: February 21, 1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plant, Thomas, G.
 ; REGISTRATION NUMBER: 35,784
 ; REFERENCE/DOCKET NUMBER: X-8231
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-276-2459
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3729 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-804-227C-4

Query Match 46.2%; Score 36; DB 2; Length 3729;
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRSALGSDGV 13
 :||| |||
 Db 1778 RSRAPAGADGV 1789

Search completed: July 7, 2003, 14:25:41
 Job time : 2.5434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:21:30 ; Search time 0.664151 Seconds
(without alignments)
2770.284 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDGVRYT 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PC1_NEM_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PC1_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	10 US-09-815-242-10152	Sequence 10152, A
2	49	62.8	269	10 US-09-815-242-13824	Sequence 13824, A
3	44	56.4	265	10 US-09-815-242-11851	Sequence 11851, A
4	41	52.6	451	9 US-10-128-714-3277	Sequence 3277, Ap
5	41	52.6	512	9 US-10-156-761-9793	Sequence 9793, Ap
6	41	52.6	559	9 US-10-128-714-8277	Sequence 8277, Ap
7	39	50.0	1000	9 US-10-156-761-13499	Sequence 13499, A
8	38	48.7	308	9 US-09-882-691-12	Sequence 12, Appl
9	38	48.7	315	10 US-09-815-242-13289	Sequence 13289, A
10	38	48.7	985	9 US-09-994-064-6	Sequence 6, Appl1
11	38	48.7	985	9 US-09-994-064-66	Sequence 66, Appl1
12	38	48.7	1233	9 US-09-738-626-4312	Sequence 4312, Ap
13	37.5	48.1	355	9 US-10-156-761-14091	Sequence 14091, A
14	37	47.4	172	9 US-10-156-761-14091	Sequence 14091, A
15	37	47.4	206	9 US-10-156-761-10243	Sequence 10243, A
16	37	47.4	320	9 US-09-880-748-1583	Sequence 1583, Ap
17	37	47.4	352	9 US-10-156-761-13166	Sequence 13166, A
18	37	47.4	338	9 US-10-156-761-13572	Sequence 13572, A
19	37	47.4	441	9 US-09-738-626-4275	Sequence 4275, Ap

20	37	47.4	511	9 US-10-156-761-8583	Sequence 8583, Ap
21	37	47.4	530	9 US-10-156-761-10819	Sequence 10819, A
22	37	47.4	587	9 US-10-156-761-8493	Sequence 8493, Ap
23	37	47.4	722	9 US-09-738-626-5453	Sequence 5453, Ap
24	37	47.4	747	9 US-10-156-761-8661	Sequence 8661, Ap
25	37	47.4	943	10 US-09-969-362-5	Sequence 5, Appl1
26	37	47.4	1225	9 US-10-156-761-12577	Sequence 12577, A
27	37	47.4	1316	9 US-10-120-544A-4	Sequence 4, Appl1
28	37	47.4	1344	9 US-10-120-544A-20	Sequence 20, Appl1
29	37	47.4	1366	9 US-10-120-544A-6	Sequence 40694, A
30	36	46.2	30	10 US-09-864-761-40694	Sequence 7, Appl1
31	36	46.2	135	9 US-10-137-077-7	Sequence 4818, Ap
32	36	46.2	187	9 US-10-106-698-4818	Sequence 11279, A
33	36	46.2	295	10 US-09-815-242-11279	Sequence 8219, A
34	36	46.2	305	9 US-10-156-761-8829	Sequence 4829, Ap
35	36	46.2	358	9 US-09-738-626-4829	Sequence 9458, Ap
36	36	46.2	377	9 US-10-156-761-9458	Sequence 11514, A
37	36	46.2	438	9 US-10-156-761-11514	Sequence 234, Ap
38	36	46.2	1160	9 US-10-028-072-234	Sequence 234, Ap
39	36	46.2	1160	9 US-10-121-049-234	Sequence 234, Ap
40	36	46.2	1160	9 US-10-123-904-234	Sequence 234, Ap
41	36	46.2	1160	9 US-10-140-470-234	Sequence 234, Ap
42	36	46.2	1160	9 US-10-175-746-234	Sequence 234, Ap
43	36	46.2	1160	9 US-10-176-918-234	Sequence 234, Ap
44	36	46.2	1160	9 US-10-176-921-234	Sequence 234, Ap
45	36	46.2	1160	9 US-10-137-865-234	Sequence 234, Ap

ALIGNMENTS

RESULT 1
US-09-815-242-10152
Sequence 10152, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 10152
LENGTH: 262
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10152
Query Match 62.8%; Score 49; DB 10; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.59;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:||||:|||||

DB 170 VRYMANMGPEGVR 184

RESULT 2

US-09-815-242-13824

Sequence 13824, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13824

LENGTH: 269

TYPE: PRT

ORGANISM: Salmonella typhi

US-09-815-242-13824

Query Match 62.8%; Score 49; DB 10; Length 269;
Best Local Similarity 60.0%; Pred. No. 0.61;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:||||:|||||

DB 177 VRYMANMGPEGVR 191

RESULT 3

US-09-815-242-11851

Sequence 11851, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11851

LENGTH: 265

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-815-242-11851

Query Match 56.4%; Score 44; DB 10; Length 265;
Best Local Similarity 53.3%; Pred. No. 4.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:||||:|||||

DB 173 VRYLNGSLGAGGTRV 187

RESULT 4

US-10-128-714-3277

Sequence 3277, Application US/10128714

Publication No. US20030119013A1

GENERAL INFORMATION:

APPLICANT: Jiang, Bo

APPLICANT: Hu, Wenqi

APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos

APPLICANT: Eroshkin, Alexey M

APPLICANT: Lemieux, Sebastien M

TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

FILE REFERENCE: 10182-018-999

CURRENT APPLICATION NUMBER: US/10/128,714

CURRENT FILING DATE: 2002-04-23

PRIOR APPLICATION NUMBER: US 60/285,697

PRIOR FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: US 60/287,066

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 60/295,890

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: US 60/303,899

PRIOR FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/316,362

PRIOR FILING DATE: 2001-08-31

NUMBER OF SEQ ID NOS: 8603

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3277

LENGTH: 451

TYPE: PRT

ORGANISM: Aspergillus fumigatus

US-10-128-714-3277

Query Match 52.6%; Score 41; DB 9; Length 451;
Best Local Similarity 60.0%; Pred. No. 31;

Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:||||:|||||

DB 14 MROEAKATGKDGPRV 28

RESULT 5

US-10-156-761-9793
; Sequence 9793, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9793
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9793

Query Match 52.6%; Score 41; DB 9; Length 512;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YRASALGSDGVR 15
DB 136 YRVYALDSKGINV 148

RESULT 6

US-10-128-714-8277
; Sequence 8277, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8277
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8277

Query Match 52.6%; Score 41; DB 9; Length 559;

Best Local Similarity 60.0%; Pred. No. 40;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
DB 138 MRQEAATGKGDGPRV 152

RESULT 7
US-10-156-761-13499
; Sequence 13499, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13499
; LENGTH: 1000
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13499

Query Match 50.0%; Score 39; DB 9; Length 1000;
Best Local Similarity 61.5%; Pred. No. 1,8e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVR 14
DB 839 RYPAEGIGDDAVR 851

RESULT 8
US-09-882-691-12
; Sequence 12, Application US/09882691
; Publication No. US20030104593A1
; GENERAL INFORMATION:
; APPLICANT: Famedu, Lavo O.
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Orozco, Buddy
; TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes
; FILE REFERENCE: BB-1176
; CURRENT APPLICATION NUMBER: US/09/882,691
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/092,952
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-882-691-12

Query Match 48.7%; Score 38; DB 9; Length 308;
Best Local Similarity 88.9%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ASALGSDGV 13
||||| |||

Db 120 ASALGSDGV 128

RESULT 9
US-09-815-242-13289; Sequence 13289, Application US/09815242
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIORITY FILING DATE: 2001-03-21

; PRIORITY APPLICATION NUMBER: 60/191,078

; PRIORITY FILING DATE: 2000-03-21

; PRIORITY APPLICATION NUMBER: 60/206,848

; PRIORITY FILING DATE: 2000-05-23

; PRIORITY APPLICATION NUMBER: 60/207,727

; PRIORITY FILING DATE: 2000-05-26

; PRIORITY APPLICATION NUMBER: 60/242,578

; PRIORITY FILING DATE: 2000-10-23

; PRIORITY APPLICATION NUMBER: 60/253,625

; PRIORITY FILING DATE: 2000-11-27

; PRIORITY APPLICATION NUMBER: 60/257,931

; PRIORITY FILING DATE: 2000-12-22

; PRIORITY APPLICATION NUMBER: 60/269,308

; PRIORITY FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13289

; LENGTH: 315

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(315)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-815-242-13289

Query Match 48.7%; Score 38; DB 10; Length 315;

Best Local Similarity 46.7%; Pred. No. 71;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 MRYASALGSDGV 15

Db 281 MRYAKEMGFDAIRL 295

RESULT 10
US-09-994-064-6

; Sequence 6, Application US/09994064

; Publication No. US20030082788A1

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/994,064

; FILING DATE:

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/468,190

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 985 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-994-064-6

Query Match 48.7%; Score 38; DB 9; Length 985;

Best Local Similarity 57.1%; Pred. No. 27+02;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 YRASALGSDGV 16

Db 921 YECTVLISDGRVT 934

RESULT 11
US-09-994-064-66

; Sequence 66, Application US/09994064

; Publication No. US20030082788A1

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/994,064

; FILING DATE:

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/468,190

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39116-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-994-064-66

Query Match 48.7% Score 38; DB 9; Length 985;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 YRASALGSDGVRV 16
DB 921 YECTVLISDGTREV 934

RESULT 12

US-09-738-626-4312
Sequence 4312, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4312
LENGTH: 1233
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4312

Query Match 48.7% Score 38; DB 9; Length 1233;
Best Local Similarity 37.5%; Pred. No. 3.6e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 16
DB 34 IQARISGIGNDGKRT 49

RESULT 13
US-10-156-761-15060
Sequence 15060, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761

CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 15060
LENGTH: 355
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-15060

Query Match 48.1% Score 37.5; DB 9; Length 355;
Best Local Similarity 56.2%; Pred. No. 1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 1 MRYRASALGSDGVRV 16
DB 1 MGRFASRIGS-GIMAT 15

RESULT 14
US-10-156-761-14091
Sequence 14091, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 14091
LENGTH: 172
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-14091

Query Match 47.4% Score 37; DB 9; Length 172;
Best Local Similarity 53.3%; Pred. No. 52;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 30 IRYPAELLHDDGTRV 44

RESULT 15
US-10-156-761-10243
Sequence 10243, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 10243
 ; LENGTH: 206
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 ; US-10-156-761-10243

Query Match 47.4%; Score 37; DB 9; Length 206;
 Best Local Similarity 53.8%; Pred. No. 65;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 RASALGSDGVRYT 16
 : | | | | : |
 Db 119 KAEALGDDYRIT 131

Search completed: July 7, 2003, 14:26:34
 Job time : 1.66415 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:12:35 ; Search time 0.815094 Seconds
(without alignments)
1887.083 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78
Sequence: 1 MRYRASALGSDGVRVT 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73:
2: PIR:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	1	enoyl-[acyl]-carrie
2	49	62.8	262	1	enoyl-[acyl]-carrie
3	49	62.8	262	2	enoyl-[acyl]-carrie
4	49	62.8	262	2	enoyl-[acyl]-carrie
5	49	62.8	262	2	enoyl-[acyl]-carrie
6	48	61.5	99	2	enoyl-[acyl]-carrie
7	45	57.7	260	2	enoyl-[acyl]-carrie
8	44	56.4	264	2	enoyl-[acyl]-carrie
9	44	56.4	265	2	enoyl-[acyl]-carrie
10	44	56.4	1112	2	enoyl-[acyl]-carrie
11	44	56.4	1385	2	enoyl-[acyl]-carrie
12	44	56.4	1391	2	enoyl-[acyl]-carrie
13	42	53.8	291	2	enoyl-[acyl]-carrie
14	41	52.6	238	2	enoyl-[acyl]-carrie
15	41	52.6	356	2	enoyl-[acyl]-carrie
16	41	52.6	415	2	enoyl-[acyl]-carrie
17	41	52.6	547	2	enoyl-[acyl]-carrie
18	41	52.6	572	2	enoyl-[acyl]-carrie
19	39	50.0	261	2	enoyl-[acyl]-carrie
20	39	50.0	261	2	enoyl-[acyl]-carrie
21	39	50.0	262	2	enoyl-[acyl]-carrie
22	39	50.0	262	2	enoyl-[acyl]-carrie
23	39	50.0	274	2	enoyl-[acyl]-carrie
24	39	50.0	274	2	enoyl-[acyl]-carrie
25	39	50.0	274	2	enoyl-[acyl]-carrie
26	38.5	49.4	1296	2	enoyl-[acyl]-carrie
27	38.5	49.4	513	2	enoyl-[acyl]-carrie
28	38	48.7	109	2	enoyl-[acyl]-carrie
29	38	48.7	146	2	enoyl-[acyl]-carrie

30	38	48.7	155	2	adenylsulfate 3-
31	38	48.7	272	2	hypothetical prote
32	38	48.7	272	2	enoyl-[acyl]-carrie
33	38	48.7	272	2	AG2669
34	38	48.7	312	2	PG5164
35	38	48.7	316	2	HPK(Ser) kinase/ph
36	38	48.7	374	2	HPK(Ser-P) kinase/p
37	38	48.7	378	1	probable hexosyltr
38	38	48.7	416	2	probable flagellar
39	38	48.7	424	2	glutamate-1-semial
40	38	48.7	437	2	hasa export system
41	38	48.7	452	2	conserved hypothet
42	38	48.7	474	1	cyae protein - Bor
43	38	48.7	492	2	hypothetical prote
44	38	48.7	513	2	indole-3-pyruvate
45	38	48.7	545	1	probable sugar tra
			547	2	

ALIGNMENTS

RESULT 1
S48029
enoyl-[acyl]-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - Escherichia coli (strain
N:Alternate names: enoyl-ACP reductase; short-chain alcohol dehydrogenase homolog env
C:Species: Escherichia coli
C>Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 03-Jun-2002
C/Accession: S48029; A47681; C64877
R/Kater, M.M.; Koningstein, G.M.; Nijkamp, H.J.J.; Stultje, A.R.
Plant Mol. Biol. 25, 771-790, 1994
A>Title: The use of a hybrid genetic system to study the functional relationship betw
A/Reference number: S48029; MUID:94355651; PMID:8075395
A/Accession: S48029
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-262 <KAT>
A/Cross-references: EMBL:X78733; NID:9587105; PIDD:CA55381.1; PID:9587106
A/Experimental source: Strain K-12, substrain W3110
R/Bergler, H.; Hogenauer, G.; Turnowsky, F.
J. Gen. Microbiol. 138, 2093-2100, 1992
A>Title: Sequences of the env gene and of two mutated alleles in Escherichia coli.
A/Reference number: A47681; MUID:93123967; PMID:1364817
A/Accession: A47681
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-262 <BER>
A/Cross-references: GB:M97219; NID:9145850; PIDD:AA17755.1; PID:9145851
A/Note: Sequence extracted from NCBI backbone (NCBI:121825, NCBI:121826)
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: C64877
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-262 <BIAN>
A/Cross-references: GB:AE00227; GB:U00096; NID:91787543; PIDD:AACT4370.1; PID:917875
A/Experimental source: Strain K-12, substrain MG1655
C/Genetics:
A/Genes: fabI, envM
C/Function:
A/Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the corre
A/Pathway: fatty acid biosynthesis
A/Note: Inhibited by palmitoyl-CoA and diazaborine
C/Superfamily: enoyl-[acyl]-carrier-protein] reductase (NADH); short-chain alcohol deh
C/Keywords: fatty acid biosynthesis; homotetramer; inner membrane; NAD; oxidoreductas
F:7-190/Domain: short-chain alcohol dehydrogenase homology <SNADH>
F:9-35/Region: NAD binding

Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:|:|:|:|:|
DB 170 VRYMANAMGPEGV 184

RESULT 2

B43729
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - *Salmonella typhimurium*
C:Species: *Salmonella typhimurium*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: B43729; S70724
R:Turnovsky, F.; Fuchs, K.; Jeschek, C.; Hoegenauer, G.
J. Bacteriol. 171, 6555-6565, 1989
A:Title: envM genes of *Salmonella typhimurium* and *Escherichia coli*.
A:Reference number: A43729; MUID:90078098; PMID:2687243
A:Accession: B43729
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <TUN>
A:Cross-references: GB:M31806; NID:g153953; PIDN:AAA27059.1; PID:g153955
R:Ol, S.Y.; Li, Y.; Sztybel, A.; Gilles, I.G.; Moltr, A.; O'Connor, C.D.
Mol. Microbiol. 17, 523-531, 1995
A:Title: *Salmonella typhimurium* responses to a bactericidal protein from human neutrophils.
A:Reference number: S70719; MUID:96100451; PMID:8559071
A:Accession: S70724
A:Molecule type: protein
A:Residues: 'X', '3', '4', 'X', '6', '11' <QIS>
A:Experimental source: strain SL1344
C:Genetics:
A:Gene: envM
C:Function:
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the correspond
A:Pathway: fatty acid biosynthesis
A>Note: inhibited by diazaboline
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehydro
C:Keywords: fatty acid biosynthesis; inner membrane; NAD; oxidoreductase
F:7-190/Domain: short-chain alcohol dehydrogenase homology <SMDH>
F:9-35/Region: NAD binding

Query Match 62.8%; Score 49; DB 1; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:|:|:|:|:|
DB 170 VRYMANAMGPEGV 184

RESULT 3

E90861
enoyl-[acyl-carrier-protein] reductase (NADH) ECS1861 [Imported] - *Escherichia coli* (str
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: E90861
R:Haysashi, T.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90861
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <NAV>
A:Cross-references: GB:BA000007; PIDN:BA035284.1; PID:g13361326; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS1861
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehyd

Query Match 62.8%; Score 49; DB 2; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:|:|:|:|:|
DB 170 VRYMANAMGPEGV 184

RESULT 4

H85757
enoyl-[acyl-carrier-protein] reductase (NADH) [Imported] - *Escherichia coli* (strain O
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85757
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
hiller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Diallantha, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <STO>
A:Cross-references: GB:AE005174; NID:g12515500; PIDN:AG56524.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: fadI
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol deh

Query Match 62.8%; Score 49; DB 2; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:~|:|:|:|:|:|
DB 170 VRYMANAMGPEGV 184

RESULT 5

AD0656
enoyl-[acyl-carrier-protein] reductase (NADH) STY1352 [Imported] - *Salmonella enteric*
C:Species: *Salmonella enterica* subsp. *enterica* serovar *typhi*
A>Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD0656
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fair
S.; Mout, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0656
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01621.1; PID:g16502475; GSPDB:GN00176
C:Genetics:
A:Gene: STY1352
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol deh

Query Match 62.8%; Score 49; DB 2; Length 262;
Best Local Similarity 60.0%; Pred. No. 0.44;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
:|||:|:~|:|:|:|:|:|
DB 170 VRYMANAMGPEGV 184

RESULT 6

S71883
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - *Proteus mirabilis* (frag
C:Species: *Proteus mirabilis*
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2002

C:Accession: S71883
R:Perito, B.; Allocati, N.; Casalone, E.; Masulli, M.; Dragani, B.; Polistinelli, M.; Acet
Biochem. J. 318, 157-162, 1996
A:Title: Molecular cloning and overexpression of a glutathione transferase gene from *Pro*
A:Reference number: S71882; MUID:96358500; PMID:8761466
A:Accession: S71883
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-99 <PER>
A:Cross-references: EMBL:U98482; NID:91053074; PIDN:AAC44361.1; PID:91053075
A:Experimental source: strain AF 2924
C:Function:
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the corresep
A:Pathway: fatty acid biosynthesis
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehyd
C:Keywords: fatty acid biosynthesis; inner membrane; NAD; oxidoreductase

Query Match 61.5%; Score 48; DB 2; Length 99;
Best Local Similarity 53.3%; Pred. No. 0.24;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 7 VRYMANMGPEGIRV 21.

RESULT 7
684960
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) [imported] - Buchnera sp. (S
C:Species: Buchnera sp.
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 03-Jun-2002
C:Accession: 684960
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: 684960
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
C:Experimental source: strain APS
C:Genetics:
A:Gene: fabI; BU265
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehyd
C:Keywords: oxidoreductase

Query Match 57.7%; Score 45; DB 2; Length 260;
Best Local Similarity 53.3%; Pred. No. 2.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 170 VRYMASLCKENIRV 184

RESULT 8
AG3354
enoyl-[acyl-carrier-protein] reductase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AG3354
R:Kamoko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium *Ana*
A:Reference number: AB1807; MUID:21599285; PMID:11759840
A:Accession: AG3354
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA876090.1; PID:917133527; GSPDB:GN00179

A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: a114391
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehy

Query Match 56.4%; Score 44; DB 2; Length 264;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 178 VRYLASLGSQNIIV 192

RESULT 9
C83419
NADH-dependent enoyl-ACP reductase PA1806 [imported] - *Pseudomonas aeruginosa* (strain
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83419
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardi, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83419
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <STO>
A:Cross-references: GB:AE004607; GB:AE004091; NID:99947789; PIDN:AAG05195.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: fabI; PA1806
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehy

Query Match 56.4%; Score 44; DB 2; Length 265;
Best Local Similarity 53.3%; Pred. No. 3.5;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
DB 173 VRYLAGSLGAEGRV 187

RESULT 10
S28289
hypothetical protein C38C10.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jan-2000
C:Accession: S28289
R:Thomas, K.
submitted to the EMBL Data Library, December 1992
A:Reference number: S28285
A:Accession: S28289
A:Molecule type: DNA
A:Residues: 1-1112 <THO>
A:Cross-references: EMBL:219153
C:Genetics:
A:Insertion: 412/1; 612/2; 670/3; 676/3; 729/3; 914/2; 984/1
C:Superfamily: *Caenorhabditis elegans* hypothetical protein C38C10.5

Query Match 56.4%; Score 44; DB 2; Length 1112;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 16
DB 1040 IRRASOMNGDVNAT 1055

RESULT 11
A88554

protein C38C10.5a [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
 C:Accession: AB8554
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: AB8554
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1385 <STO>
 A:Cross-references: GB:chr_III; PIDN:CAAY9551.1; PID:g3874875; GSPDB:GN00021; CESP:C38C10.5a
 A:Gene: C38C10.5a
 A:Map position: 3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C38C10.5

Query Match 56.4%; Score 44; DB 2; Length 1385;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRVT 16
 :|||:|:|:|
 DB 1034 IRRFASQMGDGVNAT 1049

RESULT 12
 B88554
 protein C38C10.5b [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 02-Aug-2002
 C:Accession: B88554
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: B88554
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1391 <STO>
 A:Cross-references: GB:chr_III; PIDN:CAA82365.1; PID:g3880296; GSPDB:GN00021; CESP:C38C10.5b
 A:Note: cDNA EST EMBL:CI1839 comes from this gene
 C:Genetics:
 A:Gene: C38C10.5b
 A:Map position: 3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C38C10.5

Query Match 56.4%; Score 44; DB 2; Length 1391;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRVT 16
 :|||:|:|:|
 DB 1040 IRRFASQMGDGVNAT 1055

RESULT 13
 F86451
 protein F6N18.2 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86451
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Hansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Luo, Z.A.; Luo, J.S.; Maitl, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-291 <STO>
 A:Cross-references: GB:AE005172; NID:g6714285; PIDN:AAF25981.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F6N18.2
 A:Map position: 1

Query Match 53.8%; Score 42; DB 2; Length 291;
 Best Local Similarity 66.7%; Pred. No. 8.9; 3; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDG 12
 :|||:|:|:|
 DB 130 MRYROSSLGDSG 141

RESULT 14
 D95284
 probable [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymA
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: D95284
 R:Barnett, M.D.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
 ; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
 .; Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: D95284
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-258 <KOR>
 A:Cross-references: GB:AE006469; PIDN:AAK64838.1; PID:g14523251; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Callbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
 R.; Chaill, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
 habault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma0335
 A:Genome: plasmid
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

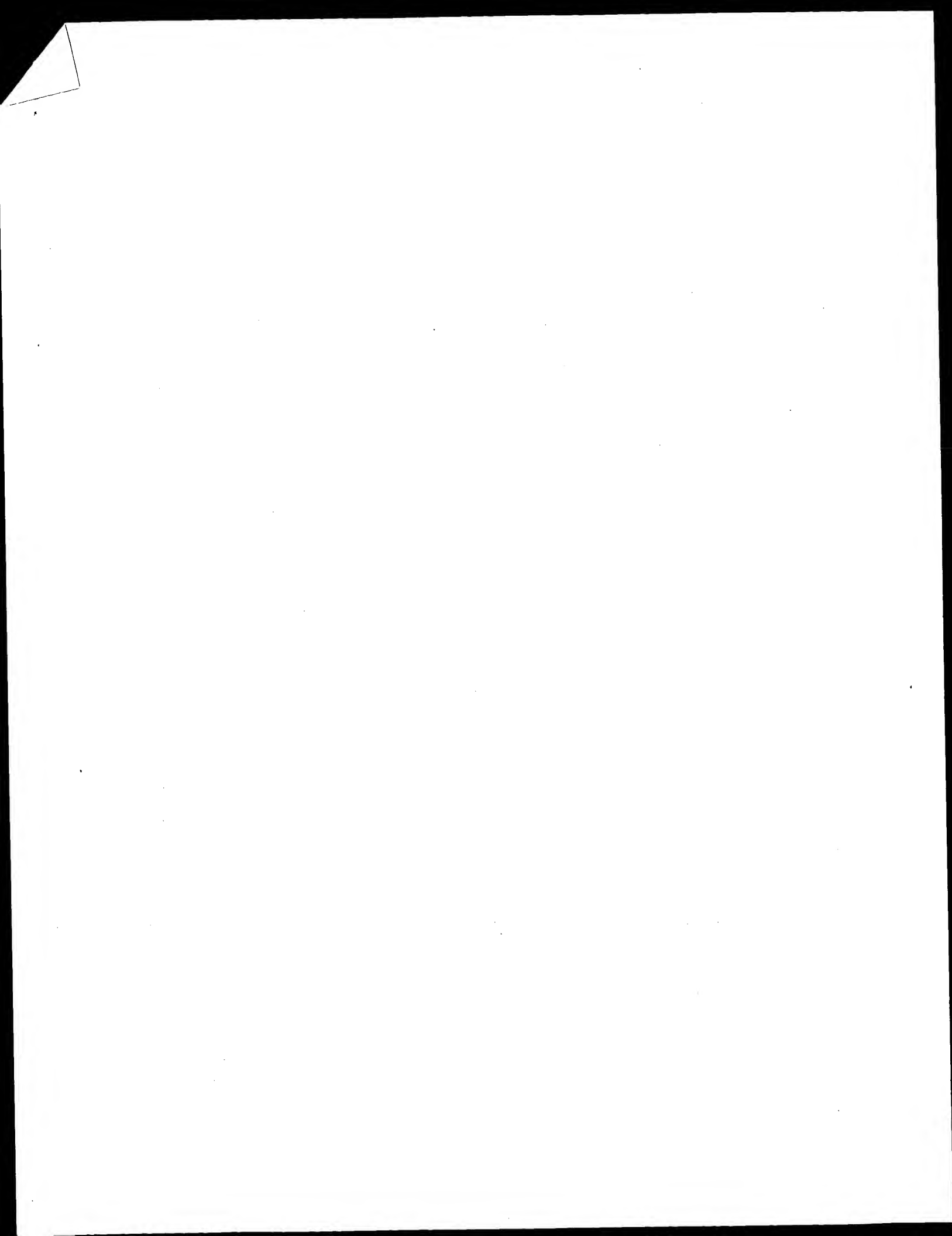
Query Match 52.6%; Score 41; DB 2; Length 258;
 Best Local Similarity 72.7%; Pred. No. 12;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 ASALGSDGVRV 15
 :|||:|:|:|
 DB 177 ADALGPDGIRV 187

RESULT 15
 T04665
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) F8D20.110 - Arabidopsi
 N:Alternate names: protein F8D20.110
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jan-2000
 C:Accession: T04665
 R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mew
 submitted to the Protein Sequence Database, July 1998
 A:Reference number: 215361

A:Accession: T04665
 A:Molecule type: DNA
 A:Residues: 1-356 <BEV>
 A:Cross-references: EMBL:AL031135; GSPDB:GN00062; ATSP:FBD20.110
 A:Experimental source: cultivar Columbia; BAC clone FBD20
 C:Genetics:
 A:Gene: ATSP:FBD20.110
 A:Map position: 4
 A:Note: Intron positions not resolved
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: phosphotransferase; protein kinase
 F:21-309/Domain: protein kinase homology <KIN>
 Query Match 52.6%; Score 41; DB 2; Length 356;
 Best Local Similarity 53.3%; Pred. No. 17;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 RYRASALGSDGVRYT 16
 DB 333 RYKAGALGAERKRRAT 347

Search completed: July 7, 2003, 14:24:55
 Job time : 1.81509 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 0.45283 Seconds

(without alignments)
1465.497 Million cell updates/sec

Title: US-09-893-371-2

Sequence: 1 MKRASALGSDGVRT 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	62.8	261	1 FABI_ECOLI	P29132 escherichia
2	49	62.8	261	1 FABI_SALTY	P16657 salmonella
3	45	57.7	260	1 FABI_BUCAI	P57353 buchnera ap
4	44	56.4	258	1 FABI_AMASP	O05069 anabaena sp
5	44	56.4	265	1 FABI_PSEAE	O92fe4 pseudomonas
6	44	56.4	1391	1 YLD5_CAEEL	O03570 caenorhabdit
7	41	52.6	548	1 AMT4_PSEST	P13507 pseudomonas
8	39	50.0	268	1 FAI2_RHIME	P58381 rhizobium m
9	38	48.7	109	1 RLAI_MAIZE	P52855 zea mays (m
10	38	48.7	155	1 CYSC_ARCFU	O29953 archaeoglob
11	38	48.7	229	1 TPIS_SULTO	O96y29 sulfobolus
12	38	48.7	272	1 FAI1_RHIME	P58380 rhizobium m
13	38	48.7	424	1 GSA_CAMTE	O99p70 campylobact
14	38	48.7	474	1 CYAE_BORE	P11092 bordetella
15	38	48.7	513	1 YH26_MZCVR	Q10510 mycobacteri
16	38	48.7	545	1 DCIP_AZOB	P15852 azospirillum
17	37	47.4	353	1 H182_MYCTU	P72039 mycobacteri
18	37	47.4	551	1 AMT4_PSESA	P23963 pseudomonas
19	37	47.4	622	1 CZAC_BACRU	O45743 bacillus th
20	37	47.4	2212	1 RRPL_EBOZM	O05318 epola virus
21	36	46.2	229	1 HIS4_PYRAE	O82y14 pyrobaculum
22	36	46.2	237	1 YPV4_METTF	P25575 methanocyst
23	36	46.2	258	1 FABI_HAEIN	P44332 synecocyst
24	36	46.2	261	1 FABI_HAEIN	P73016 rhodospirillum rubrum
25	36	46.2	334	1 BCHI_RHOSH	O30819 rhodospirillum rubrum
26	36	46.2	350	1 RPL_MICDE	P22239 rhodospirillum rubrum
27	36	46.2	361	1 DNR_STRCO	P45833 mycobacteri
28	36	46.2	401	1 DNR_STRCO	P45833 mycobacteri
29	36	46.2	427	1 TOLB_ICFPU	P44677 ictalurus p
30	36	46.2	617	1 ESRI_ICFPU	O9yhz7 ictalurus p
31	36	46.2	710	1 DAS_PICAN	P06834 picula angu
32	36	46.2	715	1 AT12_HSV1F	P08314 herpes simp
33	36	46.2	718	1 AT12_HSV1F	P10230 herpes simp

34	36	46.2	778	1 HTR6_HAIINI	O9hr92 halobacteri
35	36	46.2	778	1 HTR6_HAIINI	O48319 halobacteri
36	36	46.2	817	1 PPSA_PYRAB	P42850 pyrococcus
37	36	46.2	819	1 PPSA_PYRAB	O9v2n7 pyrococcus
38	36	46.2	821	1 PPSA_PYRAB	O57830 pyrococcus
39	36	46.2	1004	1 SPO_BACRO	P09333 bacillus br
40	36	46.2	1188	1 PPSA_METDA	O57962 methanococc
41	35	44.9	258	1 FABI_BACSU	P54616 bacillus su
42	35	44.9	285	1 G539_BACSU	P08073 bacillus su
43	35	44.9	353	1 VGLM_MCMWK	P52373 murine cyto
44	35	44.9	360	1 MTDH_ARATH	P42734 arabidopsis
45	35	44.9	368	1 RP2_STRCO	O53915 streptomyces

ALIGNMENTS

```

RESULT 1
FABI_ECOLI
ID FABI_ECOLI STANDARD: PRT: 261 AA.
AC P29132.
DT 01-DEC-1992 (Rel. 24, Created)
DR 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
EN Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
DE dependent enoyl-ACP reductase).
GN FABI OR ENVW OR B1208 OR Z2512 OR ECS1861.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-30.
RX MEDLINE=93123967; PubMed=1364817;
RA Bergler H., Hoegenauer G., Turnowsky F.;
RT "Sequences of the envm gene and of two mutated alleles in Escherichia
RT coli";
RL J. Gen. Microbiol. 138:2093-2100(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / W3110;
RA MEDLINE=94355651; PubMed=8075395;
RA Kater M.M., Koningsstein G.M., Nijkamp H.J.J., Stultje A.R.;
RT "The use of a hybrid genetic system to study the functional
RT relationship between prokaryotic and plant multi-enzyme fatty acid
RT synthetase complexes";
RL Plant Mol. Biol. 25:771-790(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1455-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RA MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Nakamura Y., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubramanian S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horikuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [5]
RP SEQUENCE FROM N.A.

```

RC STRAIN-0157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE-21074935; PubMed-11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Nanthariman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RL "Genome sequence of enterohaemorrhagic *Escherichia coli* 0157:H7,"
 Nature 409:529-533(2001).
 [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RX MEDLINE-21156321; PubMed-11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RL "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 0157:H7 and genomic comparison with a laboratory strain K-12,";
 DNA Res. 8:11-22(2001).
 [7]
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RX MEDLINE-94164884; PubMed-8119879;
 RA Bergier H., Wallner P., Ebeling A., Leitinger B., Fuchsichler S.,
 RA Aschauer H., Kollenz G., Hoegenauer G., Turnowsky F.;
 RL "Protein Erym is the NADH-dependent enoyl-ACP reductase (FabI) of
Escherichia coli,";
 J. Biol. Chem. 269:5493-5496(1994).
 [8]
 RP SEQUENCE OF 1-12.
 RC STRAIN-K12 / EMG2;
 RX MEDLINE-97443975; PubMed-9298646;
 RA Link A.J., Robinson K., Church G.M.;
 RL "Comparing the predicted and observed properties of proteins encoded
 in the genome of *Escherichia coli* K-12,";
 Electrophoresis 18:1259-1313(1997).
 [9]
 RP SEQUENCE OF 1-11.
 RC STRAIN-K12 / W3110;
 RA Frutiger S., Hughes G.J., Pasquali C., Hochstrasser D.F.;
 RL Submitted (FEB-1996) to the SWISS-PROT data bank.
 [10]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE-97113207; PubMed-8953047;
 RA Baldoek C., Rafferty J.B., Sedelnikova S.E., Baker P.J., Stuitje A.R.,
 RA Slabas A.R., Hawkes R.R., Rice D.W.;
 RL "A mechanism of drug action revealed by structural studies of enoyl
 reductase,";
 Science 274:2107-2110(1996).
 [11]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE-99425142; PubMed-10493822;
 RA Ward W.H., Holdgate G.A., Rowsell S., McLean E.G., Paupit R.A.,
 RA Clayton E., Nichols W.W., Collis J.G., Minshull C.A., Jude D.A.,
 RA Mistry A., Timms D., Camble R., Hales N.J., Britton C.J.,
 RA Taylor I.W.;
 RL "Kinetic and structural characteristics of the inhibition of enoyl
 (acyl carrier protein) reductase by triclosan,";
 Biochemistry 38:12514-12525(1999).
 [12]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RX MEDLINE-99329134; PubMed-10398587;
 RA Stewart M.J., Parikh S., Xiao G., Tonge P.J., Kisker C.;
 RL "Structural basis and mechanism of enoyl reductase inhibition by
 triclosan,";
 J. Mol. Biol. 290:859-865(1999).
 [13]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE-99215552; PubMed-10201369;
 RA Levy C.W., Roujeinikova A., Sedelnikova S., Baker P.J., Stuitje A.R.,
 RA Slabas A.R., Rice D.W., Rafferty J.B.;
 RL "Molecular basis of triclosan activity,";

RL Nature 398:383-384(1999).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
 CC -1- MISCELLANEOUS: THE ANTIBIOTIC DIAZABORINE INTERFERES WITH THE
 CC ACTIVITY BY BINDING TO THE PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M97219; AAA17755.1; -;
 CC EMBL: X78733; CA55381.1; -;
 CC EMBL: AE000227; AAC74370.1; -;
 CC EMBL: D90766; BA114841.1; -;
 CC EMBL: D90767; BA114849.1; -;
 CC EMBL: AE005378; AAC56524.1; -;
 CC EMBL: AP002556; BAB35284.1; -;
 CC PIR: A47681; A47681.
 CC PDB: 1DFG; 28-JAN-98.
 CC PDB: 1DFH; 28-JAN-98.
 CC PDB: 1DFI; 28-JAN-98.
 CC PDB: 1QG6; 21-SEP-99.
 CC PDB: 1QSG; 21-JUL-99.
 CC PDB: 1DBA; 28-OCT-99.
 CC SWISS-2DPAGE: P29132; COLI.
 CC Ecocore: BG11528; fabI.
 CC InterPro: IPR002198; ADH_short.
 CC Pfam: PF00106; adh_short; 1.
 CC Oxidoreductase; NAD; Fatty acid biosynthesis; Antibiotic resistance;
 CC Inner membrane; 3D-structure; Complete proteome.
 CC INIT MET 0 0
 CC NP_BIND 9 35 NAD (BY SIMILARITY).
 CC MUTAGEN 92 92 G->S: DIAZABORINE RESISTANCE.
 CC MUTAGEN 240 240 S->R: PRODUCES TEMPERATURE-SENSITIVE
 CC PHENOTYPE.
 CC SQ SEQUENCE 261 AA; 27732 MW; 434B019F34855956 CRC64;
 CC Query Match 62.8%; Score 49; DB 1; Length 261;
 CC Best Local Similarity 60.0%; Pred. No. 0.29;
 CC Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 1 MRYRASALGSDGVY 15
 CC :|||:|:|:|
 CC DB 169 VRYMANMGPQGVY 183
 CC
 CC RESULT 2
 CC ID: FABI_SALTY STANDARD; PRT; 261 AA.
 CC AC P16657;
 CC DT 01-AUG-1990 (Rel. 15, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 CC dependent enoyl-ACP reductase).
 CC GN FABI OR ENYM OR STM1700.
 CC OS Salmonella typhimurium
 CC CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC CC Salmonella.
 CC OX NCBI_TaxID=602;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-AG701;
 CC RX MEDLINE-90078098; PubMed-2687243;

RA Turnowsky F., Fuchs K., Jeschek C., Hoegenauer G.;
 RT "envm genes of *Salmonella typhimurium* and *Escherichia coli*.";
 RL J. Bacteriol. 171:6555-6565(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; Pubmed-11677609;
 RA McCalland M., Sanderson K.E., Spleth J., Clifton S.W., Latrelle P.,
 RA Courtney L., Portolillo S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RL L72.";
 CC Nature 413:852-856(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.
 CC -1- MISCELLANEOUS: THE ANTIBIOTIC DIAZABORINE INTERFERES WITH THE
 CC ACTIVITY BY BINDING TO THE PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M31806; AAA27059.1;
 DR EMBL: AE008775; AAL20618.1;
 DR PIR: B43729; B43729.
 DR HSP: P29132; IDPI.
 DR Styene, SG10095; fadi.
 DR Interpro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Antibiotic resistance;
 KW Inner membrane; Complete proteome.
 FT INIT MET 0 BY SIMILARITY.
 FT NP BIND 9 35 NAD (BY SIMILARITY).
 FT MUTAGEN 92 92 G->S: DIAZABORINE RESISTANCE.
 FT SEQUENCE 261 AA: 27629 MW; 187B91AE341B773 CRC64;
 SQ
 Query Match 62.8%; Score 49; DB 1; Length 261;
 Best Local Similarity 60.0%; Pred. No. 0.29;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVR 15
 DB 169 VRYMANMGPCGVR 183
 RESULT 3
 FABI_BUCAT STANDARD: PRT; 260 AA.
 AC P57353;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase).
 GN FABI OR B0265.
 OS *Buchnera aphidicola* (subsp. *Acyrtosiphon pisum*) (*Acyrtosiphon pisum*
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 CX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tokyo 1998;
 RX MEDLINE-20445173; Pubmed-10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RL *Buchnera* sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP001118; BAB12975.1;
 DR HSP: P29132; IDPI.
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Complete proteome.
 FT NP BIND 10 36 NAD (BY SIMILARITY).
 FT SEQUENCE 260 AA: 28779 MW; DCA04AC740D6DADD CRC64;
 SQ
 Query Match 57.7%; Score 45; DB 1; Length 260;
 Best Local Similarity 53.3%; Pred. No. 1.4;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OY 1 MRYRASALGSDGVR 15
 DB 170 VRYMANMGPCGVR 184
 RESULT 4
 FABI_ANASP STANDARD: PRT; 258 AA.
 AC Q05069;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase).
 GN FABI OR AL4391.
 OS *Anabaena* sp. (strain PCC 7120).
 CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 CX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-93308081; Pubmed-8391534;
 RA Wei T.-F., Ramasubramanian T.S., Pu F., Golden J.W.;
 RT "Anabaena sp. strain PCC 7120 b1fa gene encoding a sequence-specific
 RT DNA-binding protein cloned by in vivo transcriptional interference
 RT selection.";
 RL J. Bacteriol. 175:4025-4035(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-21595285; Pubmed-11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kunitz T., Sasamoto S.,
 RA Watanabe A., Iritaguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida T., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

CC -----

DR EMBL: L10036; AAD04184.1; ALT_INIT.
 DR EMBL: AP003596; BAB76090.1; ALT_INIT.
 DR HSSP: P29132; 1DPI.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR Oxidoreductase; NAD; Fatty acid biosynthesis; Complete proteome.
 KW NP_BIND 10 NAD (BY SIMILARITY).
 FT SEQUENCE 258 AA; 56.4% Score 44; DB 1; Length 258;
 SO

Query Match 56.4%; Score 44; DB 1; Length 258;
 Best Local Similarity 60.0%; Pred. No. 2.1;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 15
 DB 172 VRYLASELGSQNTIR 186

RESULT 5
 FABI_PSEAE STANDARD; PRT; 265 AA.
 ID FABI_PSEAE
 AC Q9ZFE4;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase).
 GN FABI OR PA1806.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NC NCBL_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=99395061; PubMed=10464225;
 RA Hoang T.T., Schweizer H.P.;
 RT "Characterization of Pseudomonas aeruginosa enoyl-acyl carrier protein
 RT reductase (Fabi): a target for the antimicrobial triclosan and its
 RT role in acylated homoserine lactone synthesis.";
 RT J. Bacteriol. 181:5489-5497(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gardner R.L., Goiltry L., Tolentino E., Westbrook-Wadman S., Yan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL

CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch)

CC -----

DR EMBL: AF104262; AAC95362.1; -
 DR EMBL: AE004607; AAG05195.1; -
 DR HSSP: P29132; 1DPI.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;
 KW NP_BIND 10 NAD (BY SIMILARITY).
 FT SEQUENCE 265 AA; 56.4% Score 44; DB 1; Length 265;
 SO

Query Match 56.4%; Score 44; DB 1; Length 265;
 Best Local Similarity 53.3%; Pred. No. 2.2;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 15
 DB 173 VRYLAGSLGAGTGV 187

RESULT 6
 YLDS_CAEEL STANDARD; PRT; 1391 AA.
 ID YLDS_CAEEL
 AC Q03570; P39218; P34581;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C38C10.5 in chromosome III.
 GN C38C10.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NC NCBL_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol NZ;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten N., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RT Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS, AND ALTERNATIVE SPLICING.
 RA Jones S.J.M.;
 RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a and b (shown here); may be
 CC produced by alternative splicing.
 CC -1- SIMILARITY: TO YEAST RGR1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

CC -----

DR EMBL: Z29115; CAA82365.1; JOINED.
 DR EMBL: Z19153; CAA79550.1; -
 DR EMBL: Z29115; CAA79550.1; JOINED.
 DR EMBL: Z29115; CAA82366.1; -

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Batloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 Bostard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 Godie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 Renard C., Thebault P., Vandenbol M., Weidner S., Gilbert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 Rhizobium meliloti strain 1021.";
 RL Sinchikobium meliloti strain 1021.";
 RU Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-(acyl-carrier protein) + NAD(+) = trans-
 2,3-dehydroacyl-(acyl-carrier protein) + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (by similarity).
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AL591783; CAC41683.1; -;
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;
 KM Complete proteome.
 FT NP_BIND 11 37 NAD (BY SIMILARITY).
 SQ SEQUENCE 268 AA; 28671 MW; 83D2127C7D76628C CRC64;
 QY Query Match 50.0%; Score 39; DB 1; Length 268;
 Best Local Similarity 53.3%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Db 1 MRYRASALGSDGVRV 15
 :|||: |||:|
 170 VRYLADYSGRGIRV 184
 RESULT 9
 ID RLAI_MAIZE STANDARD; PRT; 109 AA.
 AC P52855; O24414;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 60S acidic ribosomal protein P1 (L12).
 GN RPL1A.
 OS Zea mays (Maize).
 CC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 CC Panicoideae; Andropogoneae; Zea.
 CC NCBI_TaxID=4577;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CV. W22;
 RC Hamilton D.A., Turcich M.P., Bokhari-Riza A., Mascarenhas J.P.,
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 18-21.
 RC STRAIN=CV. B73; TISSUE=Ear;
 RX MBDLINE=97422884; PubMed=9276949;
 RA Bailey-Serres J., Vangala S., Szick K., Lee C.H.,
 RT "Acidic phosphoprotein complex of the 60S ribosomal subunit of maize
 seedling roots. Components and changes in response to flooding.";
 RL Plant Physiol. 114:1293-1305(1997).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
 PROTEIN SYNTHESIS (BY SIMILARITY).

CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
 SUBUNIT (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U040147; AAA91168.1; -;
 DR EMBL: U62752; AAB71079.1; -;
 DR MaltEDB: 84941; -;
 DR InterPro: IPR001813; 60S-ribosomal.
 DR Pfam: PF00428; 60S-ribosomal.1.
 KW Ribosomal protein.
 FT CONFLICT 9 R -> T (IN REF. 2).
 SQ SEQUENCE 109 AA; 11096 MW; 1BFEC8E34391F080 CRC64;
 QY Query Match 48.7%; Score 38; DB 1; Length 109;
 Best Local Similarity 46.7%; Pred. No. 9.1;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 Db 2 RYRASALGSDGVRV 16
 :|||: |||:|
 9 RYALILSDGIAIT 23
 RESULT 10
 ID CYSC_ARCFU STANDARD; PRT; 155 AA.
 AC Q29533;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
 phosphotransferase).
 DE CYSC OR AF0286.
 GN Archaeoglobus fulgidus.
 OS Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 CC Archaeoglobaceae; Archaeoglobus.
 CC NCBI_TaxID=2234;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RC MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerklavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reisch C.I., McNeill L.K., Badger J.H., Glodok A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.,
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 phosphoadenylylsulfate.
 CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. INVOLVED
 IN THE ENERGY-GENERATING SULFUR OXIDATION PATHWAYS OF MANY
 CHEMOAUTOTROPHIC MICROBES.
 CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE001085; AAB90945.1; -
 DR TIGR: AF0288; -
 DR InterPro: IPR002891; APS_kinase.
 DR Pfam: PF01583; APS_kinase; 1.
 DR ProDom: PD002350; APS_kinase; 1.
 DR TIGRfams: TIGR00455; apsk; 1.
 DR Transferrase: Kinase; ATP-binding; Phosphorylation; Complete proteome.
 FT NP_BIND 9 16 ATP (BY SIMILARITY).
 FT ACT_SITE 83 83 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 SO SEQUENCE 155 AA; 17883 MW; 4535298F48931881 CRC64;
 SIMILARITY).

Query Match 48.7%; Score 38; DB 1; Length 155;
 Best Local Similarity 57.1%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 14
 Db 29 MGYRVELDGDGVR 42

RESULT 11
 TPIS_SULTO
 ID TPIS_SULTO STANDARD; PRT; 229 AA.
 AC G96Y29;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Tricosephosphate isomerase (EC 5.3.1.1) (TIM).
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545; 7;
 RX MEDLINE=21456156; PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankel A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermocacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate -> glyceralone
 CC phosphate.
 CC -1- PATHWAY: Plays an important role in several metabolic pathways.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AP000988; BAB67127.1; -
 DR InterPro: IPR003009; FmL_enzyme.
 DR InterPro: IPR000652; Triophos_ismrse.
 DR ProDom: PD001005; Triophos_ismrse; 1.

DR TIGRfams: TIGR00419; tim; 1.
 DR PROSITE: PS00171; TIM; FALSE NEG.
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
 KW Pentose shunt; Complete proteome.
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT SEQUENCE 229 AA; 24898 MW; 9DC1D550A0183525 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 229;
 Best Local Similarity 53.8%; Pred. No. 20;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YRASALGSDGVR 15
 Db 188 YRAIELGADGIGV 200

RESULT 12
 FAIL_RHIME
 ID FAIL_RHIME STANDARD; PRT; 272 AA.
 AC P58360;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enoyl-(acyl-carrier-protein) reductase [NADH] 1 (EC 1.3.1.9) (NADH-
 DE dependent enoyl-ACP reductase 1).
 GN FABI1 OR R00898 OR SMC00005.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothé G., Ampe F., Batut J.,
 RA Bolstad P., Becker A., Boutry M., Cadiau E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 RA Pohl T., Portetle D., Puehler A., Purnelle B., Rameberger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -1- CATALYTIC ACTIVITY: Acyl-(acyl-carrier protein) + NADH.
 CC 2,3-dehydroacyl-(acyl-carrier protein) + NADH.
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AL591785; CAC45470.1; -
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;
 KW Complete proteome.
 FT NP_BIND 14 40 NAD (BY SIMILARITY).
 FT SEQUENCE 272 AA; 29148 MW; 1E0FA18A22CDBE36 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 272;
 Best Local Similarity 53.3%; Pred. No. 24;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVR 15
 Db 173 VRYIAVLGNRGIRV 187

```

RESULT 13
GSA_CAMJF
ID GSA_CAMJF STANDARD: PRT: 424 AA.
AC 09PT0;
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) (GSA)
DE (Glutamate-1-semialdehyde aminotransferase) (GSA-AT).
GN Heml OR CJO0853.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98295987; PubMed=10688204;
RA Parthill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Felwell T., Holroyd S.,
RA Jagers K., Rajashev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Raftery M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -1- CARBOLYTIC ACTIVITY: (S)-4-amino-5-oxopentanoate - 5-
CC aminelevalinate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; second step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL139076; CAB73118.1; -.
DR HSSP: P24630; 2GSA.
DR InterPro: IPR000954; Aminoctran_3.
DR InterPro: IPR004639; Heml.
DR Pfam: PF00202; aminotran_3; 1.
DR TRIGRAMS: TRIGR00713; heml; 1.
DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; FALSE_NEG.
KM Porphyrin biosynthesis; Isomerase; Pyridoxal phosphate;
KW Complete proteome. 263 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT BINDING 263
FT SQUENCE 424 AA; 46092 MW; 781152645ADDA57D CRC64;
SQ
Query Match 48.7%; Score 38; DB 1; Length 424;
Best Local Similarity 72.7%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

GN CYAE.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98295987; PubMed=2905265;
RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-
RT haemolysin bifunctional protein of Bordetella pertussis."
RL EMBL J. 7:3997-4004(1988).
CC -1- FUNCTION: CYAE IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE
CC ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRTF FAMILY OF SECRETION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14199; CAA32414.1; -.
DR PIR: S02388; BVBRCE.
DR InterPro: IPR003423; OEP.
DR Pfam: PF02321; OEP; 2.
KM Hemolysis; Transport; Outer membrane; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 474 PROTEIN CYAE.
FT SQUENCE 474 AA; 50204 MW; 29A4F21B377FC957 CRC64;
SQ
Query Match 48.7%; Score 38; DB 1; Length 474;
Best Local Similarity 63.6%; Pred. No. 44;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

RESULT 15
Y26_MYCTU
ID Y26_MYCTU STANDARD: PRT: 513 AA.
AC 010510;
DR 01-OCT-1996 (Rel. 34, Created)
DR 01-OCT-1996 (Rel. 34, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Y2226.
GN RV2226 OR MT2285 OR MYCT427.07.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
[2]

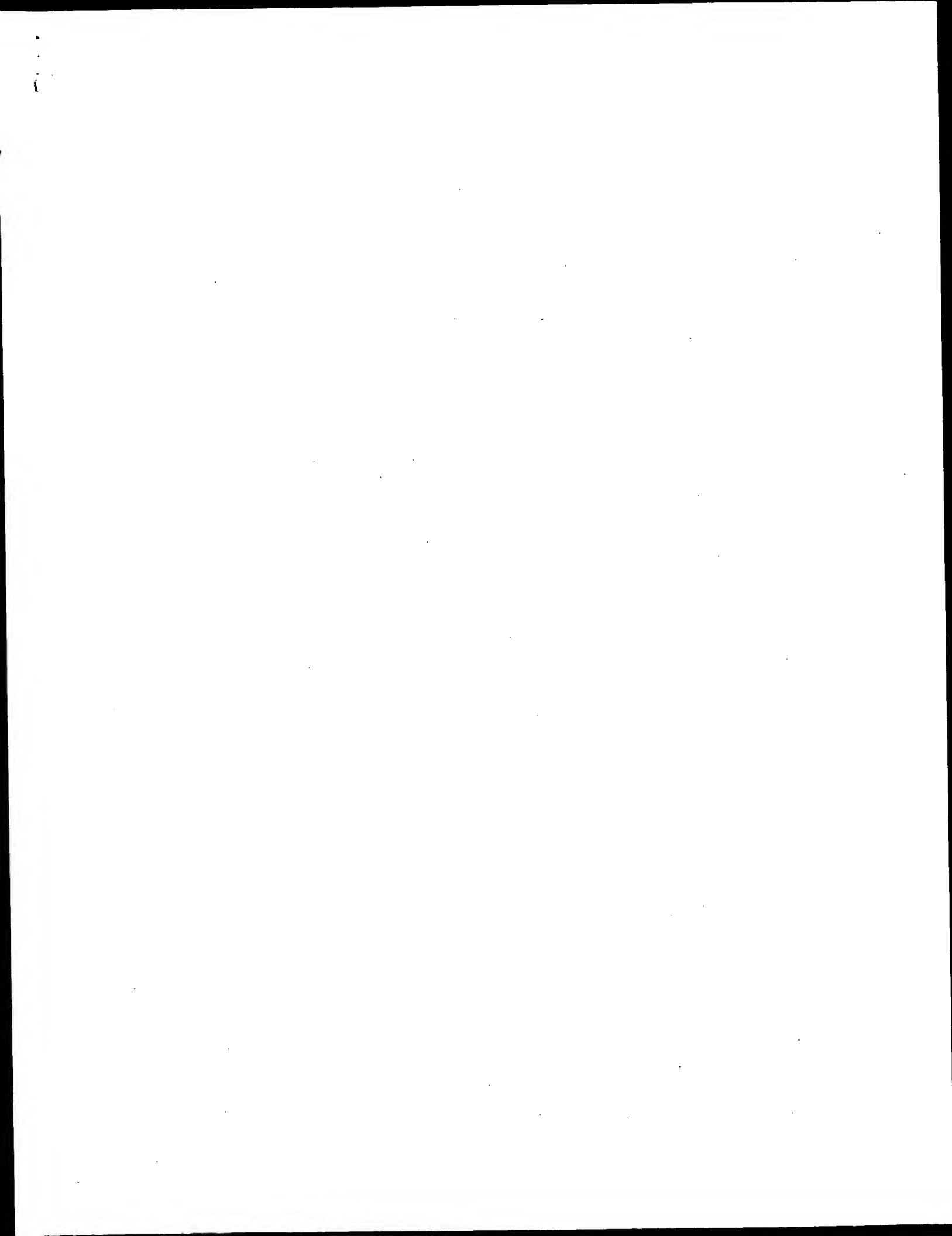
```

RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh:
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Deicher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z70692; CAA94649.1; -
 DR EMBL; AE007073; AAK46570.1; -
 DR TIGR; MT2285; -
 DR TubercuList; RV2226; -
 KM Hypothetical protein; Complete proteome.
 FT CONFLICT 299 299 D -> N (IN REF. 2).
 SO SEQUENCE 513 AA; 56333 MW; 8E74BDA8945D7B99 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 513;
 Best Local Similarity 53.8%; Pred. NO. 48;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRRASALGSDGV 13
 :|||:|:|:|
 Db 426 LRYTAATGADNV 438

Search completed: July 7, 2003, 14:21:59
 Job time : 1.45283 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2003, 14:11:44 ; Search time 1.63019 Seconds
(without alignments)
2022.314 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDGVRV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: SP-REMBL.21.*
2: SP-Dacteria.*
3: SP-Fungi.*
4: SP-Human.*
5: SP-Invertebrate.*
6: SP-Mammal.*
7: SP-Mhc.*
8: SP-Organella.*
9: SP-Phage.*
10: SP-Plant.*
11: SP-Rodent.*
12: SP-Virus.*
13: SP-Vertebrate.*
14: SP-Unclassified.*
15: SP-Virus.*
16: SP-Bacteriap.*
17: SP-Archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	1044	11 Q9ERB6	Q9ERB6 mus musculus
2	50	64.1	260	16 Q9CP73	Q9CP73 pasteurella
3	49	62.8	262	16 Q8Z7C7	Q8Z7C7 salmonella
4	48	61.5	99	2 Q51891	Q51891 proteus mir
5	45	57.7	264	16 Q8Y072	Q8Y072 raietonia s
6	45	56.4	444	5 Q8W034	Q8W034 leishmania
7	44	55.1	268	2 Q9R033	Q9R033 legionella
8	43	55.1	826	2 Q9F1K5	Q9F1K5 burkholderi
9	42	53.8	291	10 Q9LPR0	Q9LPR0 arabidopsis
10	41	52.6	151	2 Q9L5F6	Q9L5F6 salmonella
11	41	52.6	151	16 Q935M8	Q935M8 salmonella
12	41	52.6	252	16 Q92M00	Q92M00 rhizobium m
13	41	52.6	258	16 Q930L5	Q930L5 rhizobium m
14	41	52.6	356	10 Q81792	Q81792 arabidopsis
15	41	52.6	415	10 Q9M068	Q9M068 arabidopsis
16	41	52.6	419	10 Q8VZ07	Q8VZ07 arabidopsis

17	41	52.6	572	16 Q9S1S8	Q9S1S8 streptomyc
18	40	51.3	375	16 Q9F2K4	Q9F2K4 streptomyc
19	40	51.3	412	4 Q96PM2	Q96PM2 homo sapien
20	40	51.3	437	16 Q9KYZ2	Q9KYZ2 streptomyc
21	40	51.3	687	5 Q9GXY9	Q9GXY9 armligeres s
22	39	50.0	242	2 Q9L8E7	Q9L8E7 vibrio harv
23	39	50.0	261	16 Q9K1S1	Q9K1S1 neisseria m
24	39	50.0	261	16 Q9JSS8	Q9JSS8 neisseria m
25	39	50.0	262	16 Q92D51	Q92D51 listeria in
26	39	50.0	262	16 Q9Y0D5	Q9Y0D5 listeria in
27	39	50.0	274	16 Q8Y0D5	Q8Y0D5 listeria mo
28	39	50.0	467	10 Q8S1Q8	Q8S1Q8 oryza sativ
29	39	50.0	551	2 Q8R8Z3	Q8R8Z3 anabaena sp
30	39	50.0	793	2 Q9A8S5	Q9A8S5 streptomyc
31	39	50.0	793	2 Q9A8S5	Q9A8S5 streptomyc
32	39	50.0	844	10 Q94BN8	Q94BN8 lactobacill
33	39	50.0	875	10 Q9FLQ1	Q9FLQ1 arabidopsis
34	39	50.0	924	16 P73172	Q9F1L1 arabidopsis
35	39	50.0	1296	16 Q9P9W1	P73172 synecocyst
36	39	50.0	1470	5 Q9G9Y3	Q9P9W1 xylella fas
37	39	50.0	2393	2 Q9Z5F4	Q9G9Y3 leishmania
38	38.5	49.4	513	16 Q8X519	Q9Z5F4 myxococcus
39	38	48.7	110	16 Q92SV3	Q8X519 escherichia
40	38	48.7	146	10 Q64874	Q92SV3 rhizobium m
41	38	48.7	192	5 Q8T592	Q64874 arabidopsis
42	38	48.7	254	16 Q92P88	Q8T592 giardia lam
43	38	48.7	272	16 Q8UHC5	Q92P88 rhizobium m
44	38	48.7	298	10 Q9FJ82	Q8UHC5 agrobacteri
45	38	48.7	312	16 Q97Q18	Q9FJ82 arabidopsis
					Q97Q18 streptococc

ALIGNMENTS

RESULT 1

Q9ERB6 PRELIMINARY; PRT; 1044 AA.
ID Q9ERB6
AC Q9ERB6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nuclear myosin I beta.
DE MYOIC.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Peitic-Dragovich L., Stojiljkovic L., Phillimonenko A.A., Nowak G.,
RA Ke Y., Settlage R.E., Shabanowitz J., Hunt D.F., Hozak P.,
RA "A Myosin I Isoform in the Nucleus".
RT Science 0:0-0(2000).
RL EMBL; AY007255; AAC02570.1; -
DR HSSP; P08799; 1MND.
DR MGD; MGI:106612; MYOIC.
DR InterPro; IPR000048; IO.region.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR PRINTS; PR00193; MYOSTINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 2.
DR SMART; SM00242; MYSC; 1.
SQ SEQUENCE 1044 AA; 119876 MW; CDABOFF699D041C9 CRC64;

Query Match 100.0%; Score 78; DB 11; Length 1044;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRYRASALGSDGVRV 16
|||||

Db 1 MRYRASALGSDGVRV 16

RESULT 2

09CP73 PRELIMINARY; PRT; 260 AA.
 AC 09CP73;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE FABI.
 GN FABI OR PM0182.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 CC Proc Natl Acad Sci U S A. 98:3460-3465(2001).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 DR EMBL; AE006052; AK02266.1; -
 DR HSSP; P29132; I0SG.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 260 AA; 27850 MW; 922E5B8B0C8963 CRC64;

Query Match

Best Local Similarity 64.1%; Score 50; DB 16; Length 260;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVRV 15
 DB 171 RYMANALGSDGVRV 184

RESULT 3

08Z7C7 PRELIMINARY; PRT; 262 AA.
 AC 08Z7C7;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Enoyl-[acyl-carrier-protein] reductase (NADH).
 GN STY1352.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Rietveld I., Hamlin N., Hogue A., Hien T.T., Holtroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 DR EMBL; AL627270; CAD01621.1; -
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Complete proteome.

SQ SEQUENCE 262 AA; 27775 MW; 1844E8792E709A93 CRC64;

Query Match
 Best Local Similarity 62.8%; Score 49; DB 16; Length 262;
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRV 15
 DB 170 RYMANALGSDGVRV 184

RESULT 4

051891 PRELIMINARY; PRT; 99 AA.
 AC 051891;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to E.coli EnvM (Fragment).
 OS Proteus mirabilis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Proteus.
 OX NCBI_TaxID=584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF 2924;
 RX MEDLINE=93170270; PubMed=8436105;
 RA Mignogna G., Allcock N., Aceto A., Piccolomini R., Di Ilio C.,
 RA Barra D., Martini F.;
 RT "The amino acid sequence of glutathione transferase from Proteus
 RT mirabilis, a prototype of a new class of enzymes.";
 RL Eur. J. Biochem. 211:421-425(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF 2924;
 RX MEDLINE=96385500; PubMed=8761466;
 RA Perito B., Allcock N., Aceto A., Casalone E., Di Ilio C., Mesull M.,
 RA Dragani B., Polistelli M.;
 RT "Molecular cloning and overexpression of a glutathione transferase
 RT gene from Proteus mirabilis.";
 RL Biochem. J. 318:157-162(1996).
 DR EMBL; U38482; AAC44361.1; -
 DR HSSP; P29132; IDFI.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 FT NON_TER
 SQ SEQUENCE 99 AA; 10395 MW; 2E442AB04AFA031 CRC64;

Query Match
 Best Local Similarity 61.5%; Score 48; DB 2; Length 99;
 Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRV 15
 DB 7 RYMANALGSDGVRV 21

RESULT 5

08Y072 PRELIMINARY; PRT; 264 AA.
 AC 08Y072;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Probable enoyl-[acyl-carrier-protein] reductase (EC 1.3.1.9).
 GN FABI OR RSC1172 OR RS04528.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 CC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM1100;

RX MEDLINE=21681879; PubMed=11823852;
 RA Salenobat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Caltoico L.,
 RA Chendrier M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin M., Schlex T.,
 RA Siglier P., Thebault P., Whalen M., Winkler P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT Genome sequence of the plant pathogen *Ralstonia solanacearum*.
 RL Nature 415:497-502(2002).
 DR EMBL: AL646063; CAD14874.1;
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 264 AA; 28006 MW; 1CE9B4C10AE3A52 CRC64;

Query Match 57.7%; Score 45; DB 16; Length 264;
 Best Local Similarity 64.3%; Pred. No. 13;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 14
 DB 170 VRYLASALGPKGIR 183

RESULT 6

08WQ34 PRELIMINARY; PRT; 444 AA.

AC 08WQ34;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical predicted transmembrane protein P265.17, unknown
 DE function.
 GN P265.17.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Robben J., Grymonprez B., Weltjens I., Aert R., Volckaert G.,
 RA Ivens A.C., Quail M., Rajandream M.A., Barrell B.G.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome."
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL359716; CAD19425.1; -;
 KW Transmembrane.
 SQ SEQUENCE 444 AA; 49507 MW; 5E0CBA3C2868398A CRC64;

Query Match 57.7%; Score 45; DB 5; Length 444;
 Best Local Similarity 56.2%; Pred. No. 22;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 16
 DB 22 MSYSASLSSSGIEVT 37

RESULT 7

09RE03 PRELIMINARY; PRT; 268 AA.

AC 09RE03;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Enoyl reductase.
 GN FABI.

OS Legionella pneumophila.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC Legionellaceae; Legionella.
 NCBI_TaxID=446;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PHILADELPHIA 1 SGI;
 RA Kabbash C.A., Albala A., Latta P.D., Feinmark S., Weiss I., Walthe E.,
 RA Silverstein S.C., Shuman H.A.;
 RT "Antibiotic activity of gemfibrozil for L.pneumophila and
 M.tuberculosis."
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC EMBL: Y14405; CAB65183.1; -;
 DR HSSP: P29132; 10SG.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 268 AA; 28466 MW; DF6B4E16AE589E2D CRC64;

Query Match 56.4%; Score 44; DB 2; Length 268;
 Best Local Similarity 53.3%; Pred. No. 19;
 Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 15
 DB 177 VRYLASLGSRGIRI 191

RESULT 8

09FLK5 PRELIMINARY; PRT; 826 AA.

AC 09FLK5;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Chitinase.
 GN BCC1.
 OS Burkholderia cepacia (Pseudomonas cepacia).
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
 OC Burkholderia.
 NCBI_TaxID=292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matsuda Y., Kakutani K., Nonomura T., Ikeda S., Tamai T., Kuwabara Y.,
 RA Tanaka H., Toyoda H.;
 RT "DNA sequence of chitinase gene cloned from Burkholderia cepacia."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB053088; BAB20043.1; -;
 DR HSSP: O54468; 10BA.
 DR InterPro: IPR004866; CarB_Hex.
 DR InterPro: IPR001540; GH_20.
 DR Pfam: PF03173; CHB_Hex; 1.
 DR Pfam: PF00728; Glyco_Hydro_20; 1.
 DR PRINTS: PR00738; GLHYDRASE20.
 SQ SEQUENCE 826 AA; 90168 MW; E9CC7A6544DC31EC CRC64;

Query Match 55.1%; Score 43; DB 2; Length 826;
 Best Local Similarity 60.0%; Pred. No. 97;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVR 15
 DB 252 LRDRATITGLDGARV 266

RESULT 9

09LPK0 PRELIMINARY; PRT; 291 AA.

AC 09LPK0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

```

DE 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE F6N18.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids 11; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN
RN SEQUENCE FROM N.A.
RA Chao C., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Alatafi H., Bel O., Chin C., Chiu J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Huang B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thayerl A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F6N18 from chromosome
RT I."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011118; AAF25981.1; -
SQ SEQUENCE 291 AA; 32361 MW; 3F64694A4C8397EC CRC64;

Query Match
Best Local Similarity 53.8%; Score 42; DB 10; Length 291;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDG 12
DB 130 MRYRSSLGSDG 141
|||||
|

RESULT 10
OYLSF6 PRELIMINARY; PRT; 151 AA.
AC O9LSF6;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Orl, hypothetical protein.
GN R0189.
OS Salmonella typhi.
OG Plasmid R27.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN
RN SEQUENCE FROM N.A.
RA MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer."
RL Nucleic Acids Res. 28:2177-2186(2000).
DR EMBL; AF250878; AAF70026.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 151 AA; 18075 MW; 15E31FA26518CEB9 CRC64;

Query Match
Best Local Similarity 52.6%; Score 41; DB 2; Length 151;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 15
DB 107 RYRLNRIGTDGLTV 120
|||||
|

RESULT 11
OY35M8 PRELIMINARY; PRT; 151 AA.
AC O935M8;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)

```

```

DE Putative membrane protein.
GN HCM1.128.
OS Salmonella typhi.
OC Plasmid PHCM1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN
RN SEQUENCE FROM N.A.
RA STRAIN=CT18;
RA MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Kirogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL513383; CAD09726.1; -
SQ SEQUENCE 151 AA; 18158 MW; A9927F14B3C3A30F CRC64;

Query Match
Best Local Similarity 52.6%; Score 41; DB 16; Length 151;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 15
DB 107 RYRLNRIGTDGLTV 120
|||||
|

RESULT 12
OY2MO0 PRELIMINARY; PRT; 252 AA.
AC O92MO0;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Putative oxidoreductase protein (EC 1.1.1.-).
GN R02563 OR SMC02336.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN
RN SEQUENCE FROM N.A.
RA STRAIN=1021;
RA MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godtke T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purrelle B., Ransperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47142.1; -
DR InterPro: IPR002198; ADL-short.
DR Pfam: PF00106; adh_short.1.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN.1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 252 AA; 26560 MW; 0B10B7A0DBEF6628 CRC64;

Query Match
Best Local Similarity 52.6%; Score 41; DB 16; Length 252;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 15
|||||
|

```

DB 167 RHLAHAFGPDGIRV 180

RESULT 13

0930L5 PRELIMINARY: PRT: 258 AA.

AC 0930L5: PRELIMINARY: PRT: 258 AA.

DT 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Putative.

GN RA0180 OR SMA0335.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Plasmid pSyma (megaplasmid 1).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OK NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE=21396509; Pubmed=11481432;

RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,

RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,

RA Garjal M., Hong A., Hultzer L., Hyman R.W., Kahn D., Kahn M.L.,

RA Kallman S., Keating D.H., Palm C., Peck M.C., Suzzycki R., Wells D.H.,

RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.,

RT "Nucleotide sequence and predicted functions of the entire

RT Sinorhizobium meliloti pSyma megaplasmid."

RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).

RZ EMBL: AEO07211; AAK64838.1; -;

DR InterPro: IPR002198; ADH_short.

DR Pfam: PF00106; adh_short.1.

DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.

KW Plasmid; Hypothetical protein; Complete proteome.

SQ SEQUENCE 258 AA; 26146 MW; 446BABC14A2AB8B1 CRC64;

Query Match 52.6%; Score 41; DB 16; Length 258;

Best Local Similarity 72.7%; Pred. No. 59;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 ASALGSDGVRV 15

DB 177 ADALGPDGIRV 187

RESULT 14

081792 PRELIMINARY: PRT: 356 AA.

AC 081792: PRELIMINARY: PRT: 356 AA.

DT 01-NOV-1998 (TREMblrel. 08, Created)

DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Protein kinase - like protein (Fragment).

GN F8D20.110.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OK NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Koester P., Hempel S., Entian K.-D., Hohnel J., Jesse T.,

RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.,

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AL031135; CAA20030.1; -;

DR InterPro: IPR000719; Euk_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.

DR InterPro: IPR004040; STY_kinase.

DR InterPro: IPR001245; Tyr_kinase.

DR Pfam: PF00069; kinase.1.

DR PRINTS: PR00109; TYRKINASE.

DR ProDom: PD000001; Euk_kinase.1.

DR SMART: SM00221; STYK: 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW Kinase; Serine/threonine-protein kinase.

FT NON_TER 1

SQ SEQUENCE 356 AA; 40012 MW; FAD9736384FB8221 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 356;

Best Local Similarity 53.3%; Pred. No. 84;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 16

DB 333 RYKAGALGAEKRRAT 347

RESULT 15

09M068 PRELIMINARY: PRT: 415 AA.

AC 09M068: PRELIMINARY: PRT: 415 AA.

DT 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Protein kinase-like protein.

GN AT4G35600.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OK NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.,

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AL161587; CAB80276.1; -;

DR InterPro: IPR000719; Euk_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.

DR InterPro: IPR004040; STY_kinase.

DR InterPro: IPR001245; Tyr_kinase.

DR Pfam: PF00069; kinase.1.

DR PRINTS: PR00109; TYRKINASE.

DR ProDom: PD000001; Euk_kinase.1.

DR SMART: SM00221; STYK: 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 415 AA; 45879 MW; 2FAFD025A219F528 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 415;

Best Local Similarity 53.3%; Pred. No. 99;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 16

DB 392 RYKAGALGAEKRRAT 406

Search completed: July 7, 2003, 14:23:58

Job time: 4.63019 secs

